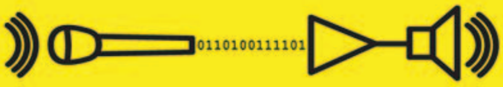
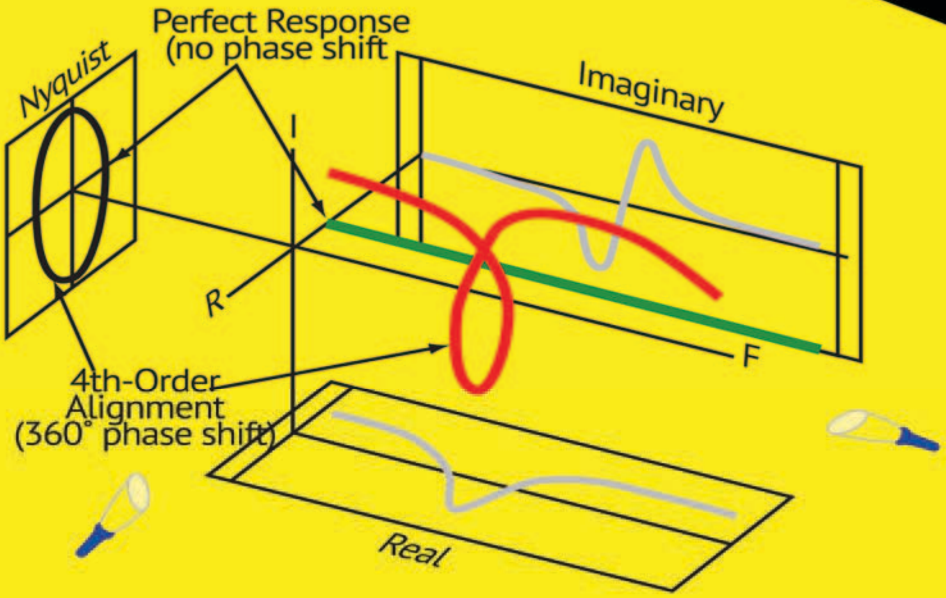
SOUND SYSTEM ENGINEERING


0110100111101

4th Edition


Perfect Response
(no phase shift)
Nyquist
Imaginary
I
R
F
4th-Order
Alignment
(360° phase shift)
Real

Don Davis

Eugene Patronis, Jr.

Pat Brown

Sound System Engineering

Sound System Engineering

Fourth Edition

Don Davis
Eugene Patronis, Jr.
Pat Brown

Edited by
Glen Ballou

Focal Press
Taylor & Francis Group
NEW YORK AND LONDON

First published 1975
by Howard W. Sams & Co., Inc.
Indianapolis, Indiana 46268

This edition published 2013
by Focal Press
70 Blanchard Road, Suite 402, Burlington, MA 01803

Simultaneously published in the UK
by Focal Press
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Focal Press is an imprint of the Taylor & Francis Group, an informa business

© 2013 Don Davis, Eugene Patronis, Jr. and Pat Brown

The right of Don Davis, Eugene Patronis, Jr. and Pat Brown to be identified as the authors of this work has been asserted by them in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging in Publication Data
CIP data has been applied for

ISBN: 978-0-240-81846-7 (hbk)
ISBN: 978-0-240-81847-4 (ebk)

Typeset in Times New Roman and Optimum
by Glen Ballou

The fourth edition of *Sound System Engineering* is dedicated to

Carolyn Davis

who is the catalyst that made it happen
and is the glue that held it all together.

Contents

Preface xv

Chapter 1 Why Sound System Engineering? 1
- Prerequisites 3
- Basic Electrical Training 3
- Mathematics 3
- Hearing Versus Listening 3
- Craftsmanship 4
- Rigging 4
- Literacy 4
- The Art, Philosophy, and Science of Sound. 4
- Fields 4

Chapter 2 Voices Out of the Past 7
- Significant Figures in the History of Audio and Acoustics 9
- 1893—The Magic Year 11
- Bell Laboratories and Western Electric 11
- Harvey Fletcher (1884–1981) 12
- Harry Nyquist (1889–1976) 12
- The dB, dBm, and the VI 12
- Sound System Equalization 13
- Acoustic Measurements—Richard C. Heyser (1931–1987) 13
- Calculators and Computers 14
- The Meaning of Communication 14
- Historical Notes 14

Chapter 3 Sound and Our Brain 17
- The Human Brain 19
- The Current Era 19
- Unexpected Validation 22

Chapter 4 Psychoacoustics 23
- Motivations 25
- Sound Reproduction 25
- Is it Better to be Born Blind or Deaf ? 26
- Recording Sound at the Eardrum 27
- Psychoacoustics via a Metaphysical Foundation 27
- Barks, Bands, Equivalent Rectangular Bandwidths (ERBs), Phons and Sones 28

Chapter 5 Digital Theory 33
- Shannon's Theory 35
- Dynamic Range 36
- The Steps from Art to Science 40
- Moravec's Warning 41
- Digital Nomenclature 42
- What Is a Bit of Data? 43
- Bayesian Theory 47
- Planck System 47
- Bits, Nats, and Bans 48
- A Communication System 48
- Holography 49

Chapter 6 Mathematics for Audio Systems 51
- Engineering Calculations 53
- Precision, Accuracy, and Resolution 54
- Simple Numbers 54
- How to Add Gains and Losses Algebraically 54
- The Factor-Label System 55

Basic Physical Terms 58
Mathematical Operations 60
Complex Number Operations 64
Decade Calibration 65
Converting Linear Scales to Logarithmic Scales 66
Finding the Renard Series for Fractional Octave Spacing 66
Radians and Steradians 67
Calculating Percentages and Ratios 70
Useful Math Tables 72
Angles 75
A Little Trigonometry 75
The Origin of the Base of the Natural Logarithm, e 76
The Complex Plane 77
Euler's Theorem 77
Examples 78
Phasors 79
Rates of Change 80

Chapter 7 Using the Decibel 85
The Decibel 87
The Neper 87
Concepts Underlying the Decibel and Its Use in Sound Systems 88
Measuring Electrical Power 90
Levels in dB 91
The Decibel in Acoustics—L_P, L_W, and L_I 92
Acoustic Intensity Level (L_I), Acoustic Power Level (L_W), and Acoustic Pressure Level (L_P) 93
Inverse Square Law 93
Directivity Factor 94
Ohm's Law 94
A Decibel Is a Decibel Is a Decibel 95
The Equivalent Level (*LEQ*) in Noise Measurements 96
Combining Decibels 97
Combining Voltage 99
Using the Log Charts 99
Finding the Logarithm of a Number to Any Base 100
Semitone Intervals 101
System Gain Changes 101
The VU and the VI Instrument 101
Calculating the Number of Decades in a Frequency Span 104
Deflection of the Eardrum at Various Sound Levels 105
The Phon 105
The Tempered Scale 106
Measuring Distortion 106
The Acoustical Meaning of Harmonic Distortion 106
Playback Systems in Studios 108
Decibels and Percentages 109
Summary 110

Chapter 8 Interfacing Electrical and Acoustic Systems 111
Alternating Current Circuits 113
Impedance 115
Electric Power 117
Properties of the LCR Circuit 120
Filters 121
Impedance Bridge 126

Constant Resistance Networks 128
Impedance Properties of Moving Coil Loudspeakers 129
Network Theorems 132
The Technician’s Viewpoint 135
Impedance Defined 135
Handling the Acoustic Input and Output of the System 137
Total Electrical Gain of a System 142
Interfacing the Electrical Output Power to the Acoustic Environment 143
Gain Structure Revisited 145
Conclusion 150

Chapter 9 Loudspeaker Directivity and Coverage 151
Essential Definitions 153
Describing Q More Accurately 158
Relationship Between $C_\angle$ and Q in an Idealized Case 159
Idealized Loudspeaker Geometry 160
Class D Audio Amplifiers 165
Sound as a Weapon 166
An Older View of Q 166
Summary 167

Chapter 10 The Acoustic Environment 169
The Acoustic Environment 171
Dispersion and Diffusion 171
Inverse Square Law 172
Atmospheric Absorption 172
Velocity of Sound 173
Isothermal vs. Adiabatic 173
Temperature-Dependent Velocity 174
The Effect of Altitude on the Velocity of Sound in Air 175
Typical Wavelengths 175
Doppler Effect 175
Reflection and Refraction 176
Effect of a Space Heater on Flutter Echo 177
Absorption 177
Definitions in Acoustics 178
Classifying Sound Fields 178
The Acoustic Environment Indoors 181
Conclusion 186

Chapter 11 Audio and Acoustic Measurements 189
Acoustic Analysis Sans Instrumentation 191
Initial Parameters 191
Acoustic Tests of Sound Systems 192
Examining AC Outlets 193
The ETC Plot 195
Site Surveys and Noise Criteria Curves 203
An Improper Use of Real Time Analysis 203
Evaluation of Listener Response 205
Fractional Bandwidth Filter Analyzers 206
Measuring Electromagnetic Pollution 209
Conclusion 211

Chapter 12 Large Room Acoustics 213
What Is a Large Room? 215
Levels Defined: Sound Power Level (L_W), Sound Intensity Level (L_I), and Sound Pressure Level (L_P) 220

Levels in Enclosed Spaces 221
Differentiating Between Reverberant Level and Reverberation Time 224
Evaluation of Signal-to-Noise Ratio, *SNR* 225
Analyzing Reflections and Their Paths 226
Critical Distance 228
Conclusion 233

Chapter 13 Small Room Acoustics 235
Non-Statistical Spaces 237
Small Room Acoustical Parameters 238
Small Room Reverberation Times 238
Small Room Resonances 239
Modes 239
What Is an Eigen Mode? 239
Small Room Geometry 240
The Initial Signal Delay Gap (*ISD*) 240
Reflections 242
Reflection Free Zone 244
Diffusion 245
Conclusion 248

Chapter 14 Designing for Acoustic Gain 251
Maximum Physical Distance 253
Establishing an Acceptable Signal-to-Noise Ratio (*SNR*) 253
Establishing an *EAD* 254
Needed Acoustic Gain (*NAG*) 254
The Number of Open Microphones 255
The Feedback Stability Margin 255
Calculating Potential Acoustic Gain 256
Obtaining ΔD_x Values 258
Measuring Acoustic Gain 259
Achieving Potential Acoustic Gain 259
Limiting Parameters in Sound Reinforcement System Design 260
How Much Electrical Power Is Required? 260
Finding the Required Electrical Power (*REP*) 261
Summary 263

Chapter 15 Designing for Speech Intelligibility 265
Introduction 267
Articulation Losses of Consonants in Speech 268
Maxfield's Equation 270
Speech Power and Articulation 270
Signal-to-Noise Ratio (*SNR*) 271
Speech Intelligibility Calculations 271
Non-Acoustic Articulation Problems 275
Relationship Between Q_{MIN} and $D_{2(MAX)}$ 276
High Density Overhead Distribution 276
$\%AL_{CONS}$ Variables 277
A Little History—Intelligibility Workshop 1986 278
Summary 279

Chapter 16 What is Waving and Why 281
General Properties of Air 283
Plane Waves 288
Non-Planar Wave Motion in a Tube 297
Plane Wave Tubes having Arbitrary Terminations 299

Impedance Tube 303
More General Waves 305
Acoustic Intensity 308
Boundaries 309
Acoustic Dipole 310

Chapter 17 Microphones 313

The Microphone as the System Input 315
Microphone Sensitivity 315
Thermal Noise 317
Microphone Selection 324
Nature of Response and Directional Characteristics 325
Boundary Microphones 330
Wireless Microphones 335
Microphone Connectors, Cables, and Phantom Power 339
Measurement Microphones 341
Microphone Calibrator 342

Chapter 18 Loudspeakers and Loudspeaker Arrays 345

Loudspeaker Types 347
Radiated Power 358
Axial Sound Pressure Level 363
Efficiency 363
Loudspeaker Electrical Impedance 364
Loudspeaker Directivity Factor 365
Loudspeaker Sensitivity 365
Direct Radiator Example Calculations 366
Horns and Compression Drivers 368
Practical Considerations Involving Horns 374
Horn Compression Drivers 376
Crossover Networks 378
Loudspeaker Arrays 392
Bessel Array 398
Line Arrays 400
Vented Enclosure Bass Loudspeakers 412
Large Signal Behavior of Loudspeakers 420

Chapter 19 Power Ratings for Amplifiers and Loudspeakers 423

Loudspeaker Power Ratings 425
Active Loudspeaker Systems 428
Non-Linear Operation 428
The Amplifier as a Voltage Source 429
The Equivalent Amplifier Size–*EAS* 430
Power from a Voltage Source 431
Burst Testing 433
Power Rating Possibilities 434
Putting It All Together 435
Multi-way Loudspeakers 438
System Gain Structure 439
Combining *MIV* and *EAS* 439

Chapter 20 Computer-Aided System Design 441

Spherical Loudspeaker Data 443
Near Field vs. Far Field 444
The Measurement Process 445
Loudspeaker Arrays 447

Direct Field Modeling 449
Room Model Detail 449
Room Acoustics—An Overview 451
Absorption 453
Realistic Room Models 457
Universal Room Modeling Tips 460
Conclusions 461

Chapter 21 Signal Delay and Signal Synchronization 463
Signal Delay 465
Useful Signal Delay Equations 466
Synchronization and Alignment of Arrays 469
Finding Acoustic Origins of Unlike Devices 470
Conclusion 472

Chapter 22 Signal Processing 475
Spectra 477
Analog to Digital Conversion 497
System Theory 502
Digital Systems and the Z Transform 520
Dynamics Processing 528

Chapter 23 Digital Audio Formats and Transports 533
The Analog Waveform 535
Quantization 535
Digital Signal Processing—DSP 542
Two Data Camps 543
How Does Ethernet Work? 545
Ethernet Protocols 546
An Open Standard 549
AES3 vs. AoE 549
Hybrid and Proprietary Systems 549
Analog vs. Digital Audio 550

Chapter 24 Sound System Equalization 553
System Criteria 555
Early Research on Equalization 555
The Transient Nature of Acoustic Feedback 556
Introduction of Real-Time Analyzers 559
Band-Rejection, Bandpass, and Band-Boost Filters 562
TEF Analysis in Equalization 567
How to Approach Equalization 568
What Can an Equalizer Equalize? 570
A Real-Time Regenerative-Response Method of Equalizing a Sound System 572
Equalizing for Playback 573
An Improper Use of Real Time Analysis in Monitoring Music and Speech 574
Diaphragmatic Absorbers 574
Don't Equalize for Hearing Loss 575
Proximity Modes 575
Checking Microphone Polarity 575
Loudspeaker Polarity 576
Summary 576

Chapter 25 Putting It All Together 577
Acoustical Analysis 579
Alternative Solutions for a Given Space 579
Device Interconnections 582

Analog Interconnection Circuitry Types 583
Signal Cables—Analog Audio, Digital Audio, and Video 590
AES3 597
Computer Control and Communication of Digital Audio 602

Appendix 605

Index 609

Preface

There are two worlds in audio—one of wave equations, Fourier, Hilbert, and Laplace transforms, and the other of Ohm's Law, Sabine and Hopkins Stryker. Eugene Patronis, Jr., straddles both like a colossus, as he is able to theorize in Quantum Mechanics and design, build, and service, with his own hands, all components used in audio.

Pat Brown is our new co-author and brings to this volume unique tools he has developed in the course of his loudspeaker testing, particularly directivity measurements, into the twenty-first century. He has taught Syn-Aud-Con seminars all over the world in person as well as through his internet training programs. He is a longtime friend of both Dr. Patronis and Don Davis, and like them, a man who delights in fully sharing his knowledge of sound system engineering with others. We welcome his participation in this volume.

The authors come from two quite different backgrounds: one is academic, the others are industrial and field oriented. *"The lion is known by his claw"* was said of Newton, whereas the technician approach uses a broad brush to get a workable, if not elegant, answer. Therefore, we have identified each author's contribution separately. It's your privilege to select the approach most applicable to your need. With today's generation of computer users and the wealth of available software it's you, the reader, who chooses the boundaries of your interests and academic skills. It is our wish that whatever background you bring to the subject you will find new tools for that level and hints of the next.

Sound System Engineering is a widely sold, widely used text on sound system design. The first editions were oriented toward those planning systems from components available in the existing marketplace, i.e., they were treated as boxes on a diagram. The first editions ignored component design and analysis other than their interconnecting parameters.

When Don and Carolyn Davis, the authors of the first two editions, sought specific advice on component design and in-depth analysis of given components they turned to their long time friend and mentor, Eugene Patronis, Jr. to provide the in-depth analysis he excels in. You will find in this edition both approaches, allowing newcomers to operate efficiently while providing the more experienced an opportunity to achieve a more advanced viewpoint. We know that one can start reading on one level, but as our experience and expertise develops, we are grateful for the more advanced approach. What we read as our learning process starts is much different years later and we become very grateful for the more advanced material.

Those who have benefited from a rigorous and thorough academic background will find that Eugene Patronis' work is a succinct summary of all you should have absorbed intellectually whereas the less sophisticated approach may contain useful nuggets that have surmounted "gray" areas in system compromises. This dual approach provides some seemingly uneven interconnects but benefits from the diverse experience of the authors.

The authors have retained their own mental images of who they are writing for, often a combination of both approaches. We hope that you will find this volume useful in pursuit of our mutual goal of truly engineered rather than merely assembled sound systems.

Thanks to Glen Ballou

Our special thanks to Glen Ballou, who transcribed our material into a publishable format. These simple words can't begin to describe the agony he has endured.

Pat Brown,

Don Davis,

Eugene Patronis, Jr.

Chapter 1

Why Sound System Engineering?

by Don Davis

1.1 Prerequisites 3
1.2 Basic Electrical Training 3
1.3 Mathematics 3
1.4 Hearing Versus Listening 3
1.5 Craftsmanship 4
1.6 Rigging 4
1.7 Literacy 4
1.8 The Art, Philosophy, and Science of Sound. 4
1.9 Fields 4

"Sound" has over the centuries been associated with human hearing (i.e.: "Is there a sound if a tree falls in the forest without a listener present?") According to Webster, "The sensation perceived by the sense of hearing." Also from Webster: "Audio, on the other hand, has largely been associated with electrical communication circuits."

"System" is a word we use to describe any "experience cluster" that we can map as a set of interacting elements over time. Typically a system is mapped by identifying the pathways of information flow, as well as possibly the flow of energy, matter, and other variables. But the flow of information is special; because only information can go from A to B while also staying at A. (Consider: photocopy machines would be useless if one didn't get to keep the original). Digital systems, analog systems, acoustic systems, etc. should be regarded by a system engineer as so many "black boxes" that need to be matched, interconnected, and adjusted. The internal circuitry should be the interest of the component designer/manufacturer.

1.1 Prerequisites

What kind of background should an aspiring sound system engineer possess is an often asked question. A list of desirable experiences would include:

1. Some basic electrical training.
2. An interest in mathematics.
3. A good ear (a love of quality sound and acute aural senses).
4. Skill with basic tools.
5. Some appreciation of the perils of rigging.
6. Good reading and writing skills.
7. A genuine appreciation for the art, philosophy, and science of sound.

1.2 Basic Electrical Training

Time spent as an apprentice electrician is not wasted. In many cases, large sound systems deal with separate power systems, and safety springs from knowledge of the power circuits that are involved. Conduits, cable sizes, and types of grounding and shielding can be complex even at power frequencies. Knowledge of the electrical codes is a necessary fundamental tool.

1.3 Mathematics

From Ohm's law to the bidding process, an ability to quickly learn new algorithms both speeds up processes and ensure profits. In today's markets "cut and try" is too expensive of both time and money to allow avoidance of basic computer skills; the use of programs such as Mathcad for both technical and financial calculations is important. Knowing what the formulae actually used are doing is essential. In order to trust any computer program, having done it first on paper the hard way, provides knowledge and confidence in the fast way and leaves you capable of detecting unexpected anomalies that might occur. Yes! You do need more than arithmetic.

1.4 Hearing Versus Listening

We all hear. But what we listen to depends to a large degree on our previous listening experiences. I have often stood in the center of an acoustic anomaly such as a reflection from an undesirable angle, distance, and level, that was destroying speech intelligibility, and watched the startled expression on the face of a person sitting in the pew as a piece of acoustical material is passed between his ears and the reflection, which restored intelligibility.

Once experienced, your eyes, ears, and brain, can recognize such problems by simply walking through them. Sensitive listening is a great plus in sound system work, and it is a sufficient reason to hear as many venues as possible under normal usage conditions. I am always surprised when I see engineers trying to design a church sound system from a set of drawings without ever having attended a service to see what their actual needs are versus what they'd like to provide them.

Because all sound system design starts in the acoustic environment and works back from there to the input, failure to experience the normal use of the space can be fatal to the ultimate end result. On one occasion I was listening in a mammoth cathedral from a position behind the altar, when asked by the administrator, if our design could solve their intelligibility problem. The priest about to conduct the service spoke to me, and because of a combination of a speech defect and in a foreign accent, I was unable to understand him to sufficiently comprehend his message. I had to tell the administrator that our system could only raise the priest's audio level, not his intelligibility.

Watching successful ministers, politicians, and other public figures use microphones reveals a world of problems unaddressed by the most competent engineer. In one case the engineer was asked if he could "put more soul in the monitor."

1.5 Craftsmanship

Possession of a guitar does not make one a musician nor do tools make a craftsman. Skill with basic tools manifests itself in clean solder joints, orderly cabling, careful labeling on panels and terminals. Construction of successful loudspeaker arrays is a challenge to both artistry and craftsmanship. In my experience craftsmanship is a direct expression of character.

1.6 Rigging

Rigging, in itself, is a business as complex and difficult as engineering the sound system and often behooves sound contractors to seek out professional assistance when required to hang large, heavy, and expensive loudspeaker arrays.

I was involved in a consulting job for a major public arena venue where the owner intended to hang the new array from the previous array's rigging. (A complicated system of cables and drums for raising and lowering the arrays). I insisted on their hiring a notable rigging authority who went up into the rigging with a camera and came down with a dozen photographs of impending disasters, such as grooves worn in the drums by the cables, frayed cables, unsafe connectors, and a lack of safety cables, to cite but a few of the problems. There are recorded fatalities from falling arrays. It is not a business for amateurs.

1.7 Literacy

This would seem obvious, but is often a weak link in an otherwise successful background experience. Sales presentations, bid offers, instruction manuals for the operators of your systems, all require reading and writing skills. Communications with customers, suppliers, and consultants needs to be thoughtfully and concisely written. For example, the contractor should be on record telling the customer that the design will function properly only if the HVAC contractor meets the specified noise criteria that is provided in the Specification. Failure to do so can be disastrous. A memo on file with the owner can save the sound contractor and/or consultant from having to take the blame.

1.8 The Art, Philosophy, and Science of Sound

The design of well-engineered sound systems stands on the shoulders of the giants who created the communication industry. "Art precedes science" is an axiom that is eternally true. Prof. Higgins as portrayed in the film, "My Fair Lady," exemplified the majesty of language, the science of studying its proper sounds, and meanings, and the engineering systems used in that earlier day. Even today the most difficult sound systems to design, build, and operate are those used in the reinforcement of live speech. Systems that are notoriously poor at speech reinforcement often pass reinforcing music with flying colors. Mega churches find that the music reproduction and reinforcement systems are often best separated into two systems

1.9 Fields

From my first view of the rainbow depiction of the electromagnetic spectrum from dc to gamma ray I have striven to gain a conceptual mental view of various fields, Fig. 1-1. Physical science, during the past century, has come to the conclusion that the Universe is some sort of field. The nature of this universal field remains controversial—is it matter which has mass? Or something more ethereal such as information?

Michael Faraday, 1831, said "*Perhaps some force is emanating from the wire.*"

A Cambridge man said "*Faraday, let me assure you, at Cambridge our electricity flows through the wire.*"

Oliver Heaviside, 1882, from his book, *Electrical Papers*, Vol. 1:

> *Had we not better give up the idea that energy is transmitted through the wire altogether? That is the plain course. The energy from the battery neither goes through the wire one way nor the other. Nor is it standing still, the transmission takes place entirely through the dielectric. What, then, is the wire? It is the sink into which the energy is poured from the dielectric and there wasted, passing from the electrical system altogether.*

John Ambrose Fleming in 1898 wrote:

> *It is important that the student should bear in mind that, although we are accustomed to speak of current as flowing through the wire in one direction or the other, this is a mere form of*

> *words. What we call the current in the wire is, to a large extent, a process going on in the space or material outside the wire….*

Ernst Guillmin, *Communications Networks*, Vol. II, 1935

> *Heaviside is the only one who considers the nature of the sources as well as the boundary effects both for the initial buildup or transient behavior and for the steady-state condition. He is the first also, to consider the leakage through the insulation, in view of which the true significance of the inductance parameter may be appreciated…. His work is a first approximation only as compared with other, more rigorous treatments. For the engineer, however, this first approximation is usually sufficient….*

Further,

> *The concept of guided waves, before Maxwell, the physical picture of the propagation of electricity through a long circuit was more or less that which is frequently presented in elementary textbooks, where the hydraulic analogy to an electric circuit is given for purposes of visualization. That is, the seat of the phenomenon was taken to be within the conductor. What occurred outside the conductor could be neither definitely formulated nor described. The electrical energy was thought of as being transmitted through the conductor which, therefore, became of prime importance. In fact, if we accept this point of view altogether, it becomes impossible to conceive of a flow of electrical energy from one point to another without the aid of an intervening conductor of some sort. It has been the writer's experience that many students are quite wedded to this point of view, so much so, in fact, that to them the propagation of energy without wires (wireless transmission) becomes a thing altogether apart from other forms of transmission involving an intervening conducting medium.*

An appreciation of Maxwell's theory of electromagnetic wave propagation brings the so-called wireless and wired forms of transmission under the same roof, so to speak. They merely appear as special cases of the same fundamental phenomenon…. The presence of a conductor merely causes the field be broken up into various components, some of which are assigned to the conductor itself, others to the surrounding medium, and still others to the surface separating the two media.

From the *Standard Handbook for Electrical Engineers* by Donald G. Fink and H. Wayne Beaty…. There is a section entitled, "Electromagnetic Wave Propagation Phenomenon."

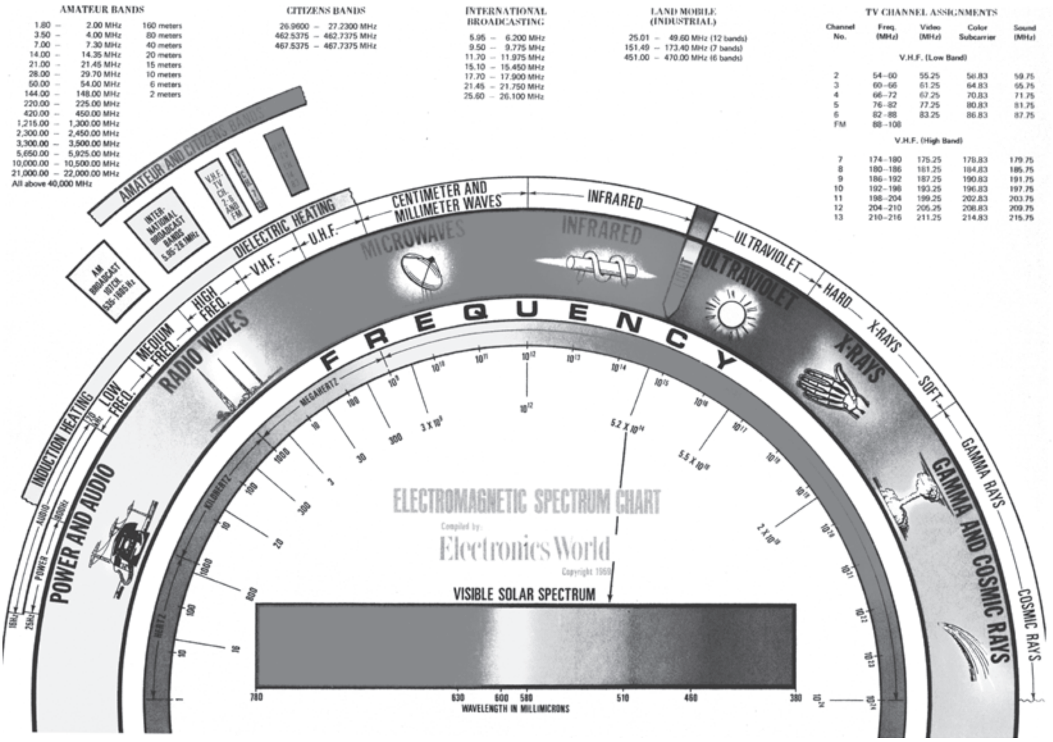
> *The usually accepted view that the conductor current produces a magnetic field surrounding it must be displaced by the more appropriate one that the electromagnetic field surrounding the conductor produces, through a small drain on the energy supply, the current in the conductor. Although the value of the latter may be used in computing transmitted energy, one should clearly recognize that physically this current produces only a loss and in no way has a direct part in the phenomenon of power transmission.*

Ralph Morrison's website has some comments on electromagnetic laws.

> *The laws I want to talk about are the basic laws of electricity. I'm not referring to circuit theory laws as described by Kirchhoff or Ohm but the laws governing the electric and magnetic fields. These fields are fundamental to all electrical activity whether the phenomenon is lightning, electrostatic display, radar, antennas, sunlight, and power generation, analog or digital circuitry. These laws are often called Maxwell's equations. Light energy can be directed by lenses, radar energy can be directed by waveguides and the energy and power frequencies can be directed by copper conductors. Thus we direct energy flow at different frequencies by using different materials. For utility power the energy travels in the space between conductors not in the conductors. In digital circuits the signals and energy travel in the spaces between traces or between the traces and the conducting surfaces. Buildings have halls and walls. People move in the halls not the walls. Circuits have traces and spaces, signals and energy moves in the spaces not in the traces.*

Scanning the Electromagnetic Spectrum Chart from dc through radio waves, light itself, out to gamma rays we can see that electromagnetic fields play a key part in our lives, Fig. 1-1.

Researchers studying human consciousness are finding electromagnetic phenomenon in addition to the previously known electrical phenomena. When EMI (electromagnetic interference) occurs in audio systems RF spectrum analyzers can be useful tools.


AMATEUR AND CITIZENS BANDS
AM BROADCAST 10TCH 535-1605 kHz
INTER-NATIONAL BROADCAST BANDS 5.95-26 MHz
V.H.F. TV CH. 2-6 AND FM
INDUCTION HEATING
DIELECTRIC HEATING
LOW FREQ.
MEDIUM FREQ.
HIGH FREQ.
V.H.F.
U.H.F.
CENTIMETER AND MILLIMETER WAVES
INFRARED
ULTRAVIOLET
HARD
X-RAYS
SOFT
GAMMA RAYS
COSMIC RAYS
POWER
AUDIO
POWER AND AUDIO
RADIO WAVES
MICROWAVES
INFRARED
LIGHT
ULTRAVIOLET
X-RAYS
GAMMA AND COSMIC RAYS
FREQUENCY
MEGAHERTZ
KILOHERTZ
HERTZ
3 X 10⁸
5.2 X 10¹⁴
5.5 X 10¹⁶
2 X 10¹⁹
ELECTROMAGNETIC SPECTRUM CHART
Compiled by
Electronics World
Copyright 1969
VISIBLE SOLAR SPECTRUM
780
630
600
580
510
460
380
WAVELENGTH IN MILLIMICRONS

AMATEUR BANDS

1.80 –	2.00 MHz	160 meters
3.50 –	4.00 MHz	80 meters
7.00 –	7.30 MHz	40 meters
14.00 –	14.35 MHz	20 meters
21.00 –	21.45 MHz	15 meters
28.00 –	29.70 MHz	10 meters
50.00 –	54.00 MHz	6 meters
144.00 –	148.00 MHz	2 meters
220.00 –	225.00 MHz	
420.00 –	450.00 MHz	
1,215.00 –	1,300.00 MHz	
2,300.00 –	2,450.00 MHz	
3,300.00 –	3,500.00 MHz	
5,650.00 –	5,925.00 MHz	
10,000.00 –	10,500.00 MHz	
21,000.00 –	22,000.00 MHz	
All above 40,000 MHz		

CITIZENS BANDS

26.9600 – 27.2300 MHz
462.5375 – 462.7375 MHz
467.5375 – 467.7375 MHz

INTERNATIONAL BROADCASTING

5.95 – 6.200 MHz
9.50 – 9.775 MHz
11.70 – 11.975 MHz
15.10 – 15.450 MHz
17.70 – 17.900 MHz
21.45 – 21.750 MHz
25.60 – 26.100 MHz

LAND MOBILE (INDUSTRIAL)

25.01 – 49.60 MHz (12 bands)
151.49 – 173.40 MHz (7 bands)
451.00 – 470.00 MHz (6 bands)

TV CHANNEL ASSIGNMENTS

Channel No.	Freq. (MHz)	Video (MHz)	Color Subcarrier	Sound (MHz)
	V.H.F. (Low Band)			
2	54–60	55.25	58.83	59.75
3	60–66	61.25	64.83	65.75
4	66–72	67.25	70.83	71.75
5	76–82	77.25	80.83	81.75
6	82–88	83.25	86.83	87.75
FM	88–108			
	V.H.F. (High Band)			
7	174–180	175.25	178.83	179.75
8	180–186	181.25	184.83	185.75
9	186–192	187.25	190.83	191.75
10	192–198	193.25	196.83	197.75
11	198–204	199.25	202.83	203.75
12	204–210	205.25	208.83	209.75
13	210–216	211.25	214.83	215.75

Figure 1-1. Electromagnetic spectrum chart.

Chapter 2

Voices Out of the Past

by Don Davis

2.1 Significant Figures in the History of Audio and Acoustics 9
2.2 1893—The Magic Year 11
2.3 Bell laboratories and Western Electric 11
2.4 Harvey Fletcher (1884–1981) 12
2.5 Harry Nyquist (1889–1976) 12
2.6 The dB, dBm, and the VI 12
2.7 Sound System Equalization 13
2.8 Acoustic Measurements—Richard C. Heyser (1931–1987) 13
2.9 Calculators and Computers 14
2.10 The Meaning of Communication 14
2.11 Historical Notes 14

During the fall of 1978, we stopped in Williamsburg, Virginia. As is our habit, we explored the old-bookstores and asked whether they had any books on acoustics. We were told they had just one, an old one. They brought out a vellum bound first edition dated 1657, *Magiae Universalis* by Gaspare P. Schotto (1608–1666).

He was a colleague of Athanasius Kircher (1602–1680) whose work was discussed in Frederick Vinton Hunt's valuable book *Origins in Acoustics*. The book, written in Latin, was published at Herbipoli, the modern Wurzburg, Germany. Some of the plates from it are shown in Figs. 2-1 through 2-6.

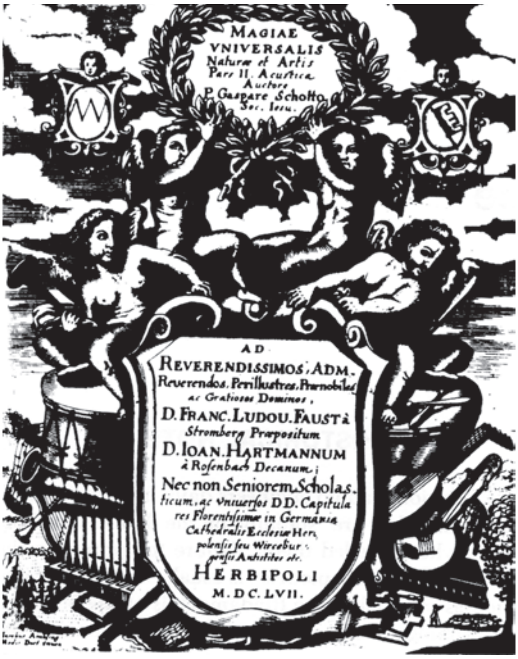

MAGIAE
VNIVERSALIS
Naturæ et Artis
Pars II. Acustica.
Auctore
P. Gaspare Schotto
Soc. Iesu.
AD
REVERENDISSIMOS, ADM.
Reverendos, Perillustres, Prænobiles
ac Gratiosos Dominos,
D. FRANC. LUDOU. FAUST à
Stromberg Præpositum
D. IOAN. HARTMANNUM
à Rosenbach Decanum;
Nec non Seniorem Scholasticum, ac universos DD. Capitulares Florentissimæ in Germania Cathedralis Ecclesiæ Herbipolensis seu Wirceburgensis Antistites etc.
HERBIPOLI
M.DC.LVII.

Figure 2-1. Frontpiece of a book published in 1857 at Herbipoli, the modern Wurzburg, Germany.

In researching the history of this book we found that it was mentioned in the Edinburgh magazine (volume 12, page 322) in 1790, as well as in Hunt's, *Origins in Acoustics*, regarding Boyle and Hooke's work. *Magiae Universalis* was used by Robert Boyle (1627–1691) and his assistant Robert Hooke (1635–1703) as they worked to improve air pumps and experiment with ticking watches in vacuums.

This book described Otto Von Guericke's work with air pumps.

An erudite discussion of Athanasius Kircher's book, *Phonurgia Nova* by Lamberto Tronchin in January 2009 edition of *Acoustics Today*, provides one of the best surveys of the period under discussion. Richard C. Heyser put it best when he said, "*You don't really own that book, you are its temporary custodian*."

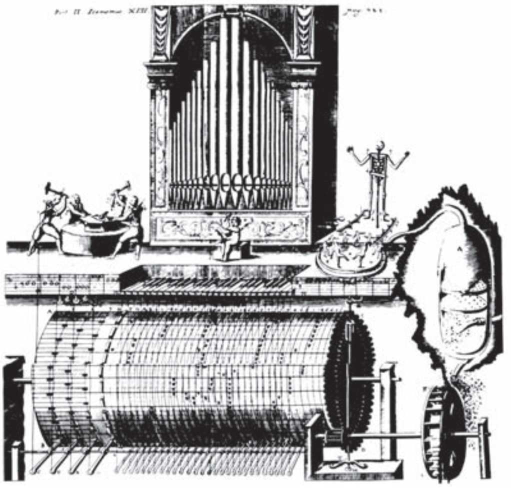


Figure 2-2. Water-powered musical instrument that fascinated our forefathers as much as computers interest us today. From *Magiae Universalis*.

These men were Galileo's contemporaries and were representative of the desire for scientific knowledge, as well as the collectors of all the myths of their age.

2.1 Significant Figures in the History of Audio and Acoustics[*]

Often, in the modern scheme of things, history is not mandatory in engineering classes taught at the university level. Significant historical figures are encountered as Faraday's ice-pail experiment, Maxwell's equations, Ohms' law etc., but not studied in depth.

Electrical engineers encounter the SI terms, potential difference in volts (Alessandro Volta), current in amperes (Andre Marie Ampere), capacitance in farads (Michael Faraday), and thermodynamic temperature in Kelvin (Lord Kelvin), as units of measurement. These were living breathing men who had occasion to interact with each other and intermingle their ideas to the benefit of science. Great seminal ideas belong to the individual, but the inter-

---

[*] Significant Figures in the History of Audio and Acoustics is an edited version of the chapter, "*Audio & Acoustic DNA—Do you Know Your Audio and Acoustic Ancestors*?" in the 4th Edition of the *Handbook for Sound Engineers*, edited by Glen Ballou.

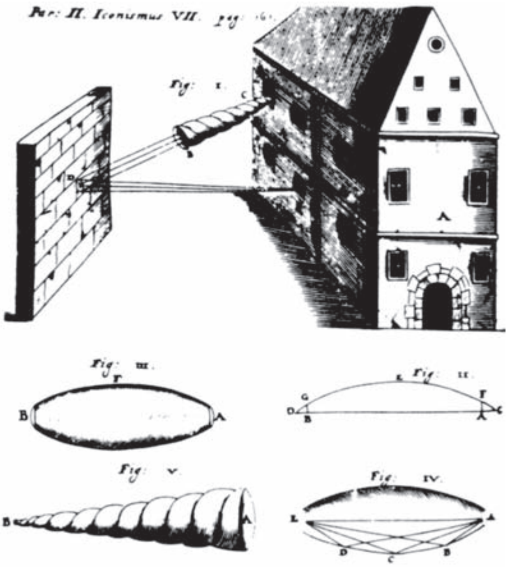

Par: II. Iconismus VII. pag: 16.
Fig: I
Fig: III
Fig: II
Fig: V
Fig: IV
A
B
C
D
E
F
G

Figure 2-3. The first bugging system. Horns such as these were used by Athanasius Kircher, a contemporary of Kasper Schott, to speak to the gatekeeper from his quarters and to eavesdrop on the conversation taking place in the courtyard. His experimental horn was 22 palms long. (A palm is about 8.7 inches, so his horn was about 16 feet long). From *Magiae Universalis*.

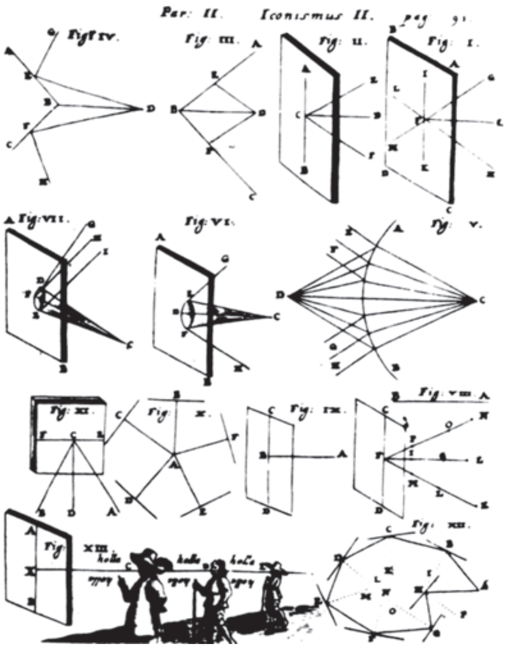

Par: II. Iconismus II. pag. 91
Fig: IV
Fig: III
Fig: II
Fig: I
Fig: VII
Fig: VI
Fig: V
Fig: XI
Fig: X
Fig: IX
Fig: VIII
Fig: XIII
Fig: XII
hollo

Figure 2-4. The basic rules for sound reflection as a geometric problem. From *Magiae Universalis*.

mingling of them leads industries. Their predecessors and contemporaries such as Joule (work, energy, heat), Charles Coulomb (electric charge), Isaac

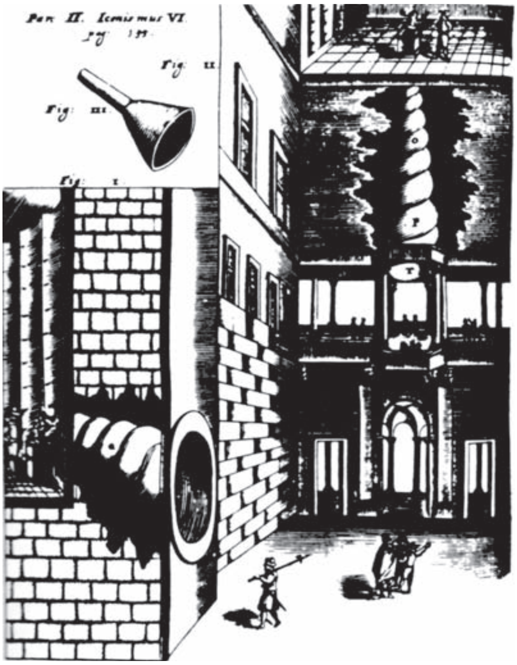

Par: II. Iconismus VI. pag: 1
Fig: II
Fig: III
Fig: I

Figure 2-5. How oracles talk or music can be transmitted from one space to another. From *Magiae Universalis*.

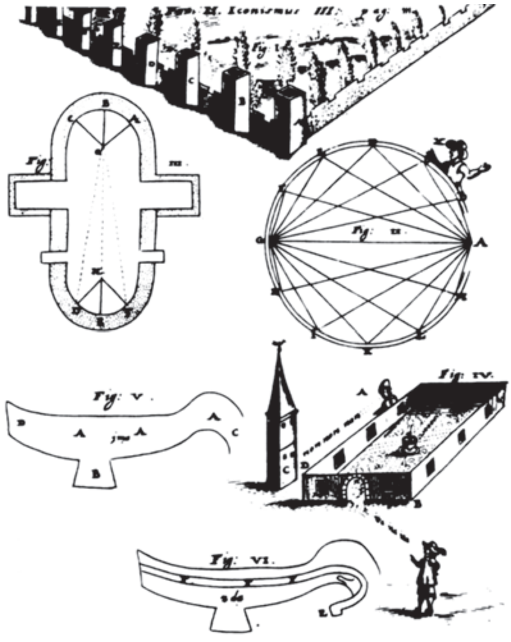

Par: II. Iconismus III. pag.
Fig: I
Fig: III
Fig: II
Fig: IV
Fig: V
Fig: VI
A
B
C

Figure 2-6. Illustrations of reflections, focusing, diffusion, time delay, and creeping. From *Magiae Universalis*.

Newton (force), Hertz (frequency), Watt (power, radiant flux), Weber (magnetic flux), Tesla (magnetic flux density), Henry (inductance), and Siemens (conductance), are immortalized as international SI derived units. Kelvin and Ampere alone have names inscribed as SI base units. Kirchhoff diagrams define the use of these units in circuit theory.

What the study of these men's lives provides, to the genuinely interested reader, is the often unique way the great ideas came to these men, the persistent pursuit of the first glimmer, and the serendipity that comes from sharing ideas with other talented minds. As all of this worked its way into the organized thinking of mankind, one of the most important innovations was the development of technical societies formed around the time of Newton, where ideas could be heard by a large receptive audience.

Some of the world's best mathematicians struggled to quantify sound in air, in enclosures and in all manner of confining pathways. Since the time of Euler (1707–1783), Lagrange (1736–1813), and d'Alembert (1717–1783), mathematical tools existed to analyze wave motion and develop field theory.

By the birth of the twentieth century, workers in the telephone industry comprised the most talented mathematicians and experimenters in both what was to become electronics and in acoustics. At MIT, the replacement of Oliver Heaviside's operational calculus by Laplace transforms gave them an enviable technical lead in education.

2.2 1893—The Magic Year

At the April 18, 1893 meeting of the American Institute of Electrical Engineers in New York City, Arthur Edwin Kennelly (1861–1939) gave a paper entitled "Impedance."

That same year General Electric, at the insistence of Edwin W. Rice, purchased Rudolph Eickemeyer's company for his transformer patents. The genius, Charles Proteus Steinmetz (1865–1923), worked for Eickemeyer. In the saga of great ideas, I have always been as intrigued by the managers of great men, as much as the great men themselves. E. W. Rice of General Electric personified true leadership when he looked past the misshaped dwarf that was Steinmetz, to the mind present in the man. General Electric's engineering preeminence, in those years, proceeded directly from Rice's extraordinary hiring of Steinmetz.

Dr. Michael I. Pupin of Columbia University was present at the Kennelly paper. Pupin mentioned Oliver Heaviside's use of the word impedance in 1887. This meeting established the correct definition of the word and established its use within the electric industry. Kennelly's paper, along with the groundwork laid by Oliver Heaviside in 1887, was instrumental in introducing the terms being established in the minds of Kennelly's peers.

The truly extraordinary Arthur Edwin Kennelly, (1861–1939) left school at the age of thirteen and taught himself physics while working as a telegrapher. He is said to "have planned and used his time with great efficiency," which is evidenced by his becoming a member of the faculty of Harvard in 1902 while also holding a joint appointment at MIT from 1913–1924. He was the author of ten books and the co-author of eighteen more, as well as writing more than 350 technical papers. Edison had employed A. E. Kennelly to provide physics and mathematics to Edison's intuition and "cut and try" experimentation. The reflecting ionosphere theory is jointly credited to Kennelly and Heaviside, and known as the Kennelly-Heaviside layer. One of Kennelly's PhD students was Vannevar Bush, who ran America's WWII scientific endeavors.

Steinmetz was not at the April 18, 1893 meeting, but sent in a letter-of-comment which included:

> *It is however, the first instance here, so far as I know, that the attention is drawn by Mr. Kennelly to the correspondence between the electrical term 'impedance' and the complex numbers. The importance hereof lies in the following: the analysis of the complex plane is very well worked out, hence by reducing the technical problems to the analysis of complex quantities they are brought within the scope of a known and well understood science.*

Nikola Tesla (1856–1943) working with Westinghouse designed the ac generator that was chosen in 1893 to power the Chicago World's Fair.

2.3 Bell laboratories and Western Electric

The University of Chicago, at the end of the Nineteenth Century and the beginning of the Twentieth Century, had Robert Millikan, America's foremost physicist. Frank Jewett, who had a doctorate in physics from MIT, and now worked for Western Electric, was able to recruit Millikan's top students. George A. Campbell (1870–1954) had by 1899 developed successful "loading coils" capable of extending the range and quality of the, at that time, unamplified telephone circuits. Unfortunately, Prof. Michael Pupin had also conceived the idea and beat him to the patent office. Bell telephone paid Pupin $435,000 for the patent and by 1925, the Campbell designed loading coils had saved Bell Telephone Company $100,000,000 in the cost of copper wire alone.

To sense the ability of loading coils to extend the range of unamplified telephone circuits, Bell had reached New York to Denver by their means alone. Until Thomas B. Doolittle evolved a method in 1877 for the manufacture of hard drawn copper, the metal had been unusable for telephony due to its inability to support its own weight over usable distances. Copper wire went from a tensile strength of 28,000 pounds per square inch with an elongation of 37%, to a tensile strength of 65,000 pounds per square inch and a elongation of 1%.

H. D. Arnold, with the advent of usable copper wire, the vacuum tube amplifier, 130,000 telephone poles, and 25 tons of copper wire was able to establish transcontinental telephony in the year 1915. There was also a public address system at those ceremonies celebrating this accomplishment.

2.4 Harvey Fletcher (1884–1981)

In 1933, Harvey Fletcher, Steinberg and Snow, Wente and Thuras and a host of other Bell Labs engineers gave birth to “Audio Perspective” demonstrations of three channel stereophonic sound capable of exceeding the dynamic range of the live orchestra. Their exhaustive study led to the conclusion that to reproduce the sound field would require an infinite number of sources and that the best compromise lay in three channels. They also demonstrated that two channels were sufficient over headphones for binaural recordings. They understood that for stereophonic reproduction each of the three channels had to cover all of the audience, a fact many contemporaries are unaware of. Early in my career at Altec, I had the privilege of working with William Snow and having full discussions of their 1933 system.

Edward C. Wente and Albert L.Thuras were responsible for the full range, low distortion, high-powered sound reproduction using condenser microphones, compression drivers, multicellular exponential horns, and horn loaded low-frequency enclosures, all of which were their original designs. The Fletcher loudspeaker, as designed by Wente and Thuras, was a three-way unit consisting of an 18 inch low-frequency driver horn loaded woofer, the incomparable W. E. 555 as a mid range, and the W. E. 597A high-frequency unit.

The power amplifiers and transmission lines were capable of full dynamic range from 30 Hz to 15,000 Hz, capabilities often claimed today but seldom realized.

In 1959, Carolyn and I repeated the original geometry tests of the 1933 experiments while working for Klipsch and Associates. Mr Klipsch and I then traveled to Bell Telephone Laboratories in New Jersey, where we made a demonstration of our results using Klipsch horns in the Arnold Auditorium. After our demonstration we were shown one of the original Fletcher loudspeakers.

A perspective can be gained, compared to today’s products, when it is realized that Western Electric components like the 555 and the 597 are to be found today in Japan, where originals sell for up to five figures. It is estimated that 99% of the existing units are in Japan. It is of interest to note that many of today’s seekers of quality sound reproduction are still building tube-type amplifiers employing the W. E. 300 B vacuum tubes.

2.5 Harry Nyquist (1889–1976)

The word inspired means “to have been touched by the hand of God.” Harry Nyquist’s thirty-seven years and 138 US patents while at Bell Telephone Laboratories personifies “inspired.” In acoustics the Nyquist plot is, by far, my favorite for a first look at an environment driven by unknown source. Nyquist also worked out the mathematics that allowed amplifier stability to be calculated leaving us the Nyquist plot which is one of the most useful audio and acoustic analysis tools ever developed. His cohort, Hendrik Bodie, gave us the frequency and phase plots as separate measurements.

Karl Kupfmuller (1897–1977) was a German engineer who paralleled Nyquist work, independently deriving fundamental results in information transmission, in closed loop modeling, including a stability criterion. Kupfmuller as early as 1928 used block diagrams to represent closed loop linear circuits. He is believed to be the first to do so.

Today’s computers, as well as digital audio devices, were first envisioned in the mid-1800s by Charles Babbage. The mathematics discussed by Lady Lovelace, the only legitimate daughter of Lord Byron, even predicted the use of a computer to generate musical tones.

Claude Shannon went from Nyquist’s paper on the mathematical limit of communication to develop “Information Theory,” which is so important to today’s communication channels.

2.6 The dB, dBm, and the VI

The development of the dB from the mile of standard cable by Bell Labs, their development and sharing of the decibel, the dBm, and the VU via the design of VI devices, changed system design into an engineering design.

The first motion pictures were silent. Fortunes were made by actors who could convey visual emotion. When motion pictures acquired sound in 1928, again via Western Electric's efforts, a large number of these well-known personalities failed to make the transition from silent to sound. The faces and figures failed to match the voices the minds of the silent movie viewers had assigned them. Later, when radio became television almost all the radio talent was able to make the transition because the familiar voices predominated over any mental visual image the radio listener had assigned to that performer. Often, at the opera, the great voices will not look the part, but just a few notes nullify any negative visual impression for the true lover of opera, whereas appearance will not compensate for a really bad voice.

In 1928, a group of Western Electric engineers became the Electrical Research Products, Inc. (ERPI), in order to service the motion picture theaters using Western Electric sound equipment. At the termination of World War II, the standard for the best bass reproduction was the loudspeakers installed in the better motion picture theaters. The goal of the designers of consumer component audio reproduction was to approach the motion picture theater quality. Western Electric had decimated their competition, RCA, in the theater business, so RCA went to court and obtained a consent decree which restricted Western Electric in the field of motion picture sound. At this point some of the engineers involved formed All Technical Services (Altec), which is why it is pronounced all tech, not al-tect. One of the pioneer engineers told me, "Those days were the equivalent of one ohm across Fort Knox." They bought the Western Electric theater inventory for pennies on the dollar.They also bought the Lansing Manufacturing Company and Peerless Manufacturing which brought James B. Lansing, Ercel Harrison and Bill Martin (Jim Lansing's brother) into Altec.

2.7 Sound System Equalization

Dr. Wayne Rudmose was the earliest researcher to perform meaningful sound system equalization. Dr. Rudmose published a truly remarkable paper in *Noise Control* (a supplementary Journal of the Acoustical Society of America) in July 1958. At the AES session in the fall of 1967, I gave the first paper on the one-third of an octave contiguous equalizer, which Altec named Acousta-Voicing. Dr. Rudmose was chairman of that AES session. The control these equalizers allowed over acoustic feedback in live sound systems quickly led to much more powerful sound reinforcement systems.

I introduced variable system equalization in special sessions at the screening facilities in August 1969 to the head sound men, Fred Wilson at MGM, Herb Taylor at Disney, and Al Green at Warner Bros. Seven Arts. These demonstrations were prior to my leaving Altec to start Synergetic Audio Concepts and others reaped the benefits of this work.

Some early workers in equalization imagined they were equalizing the room; equalization is electrical, not acoustical, and what it always adjusts is the input to the loudspeaker terminals. It allowed feedback in a reinforcement system, containing a highly efficient, but uneven amplitude response loudspeaker, to be controlled, while increasing the acoustic energy in the room.

2.8 Acoustic Measurements—Richard C. Heyser (1931–1987)

Plato said, "God ever geometrizes." Richard Heyser, the geometer, should feel at ease with God. To those whose minds respond to the visual, Heyser's measurements shed a bright light on difficult mathematical concepts. Working from Dennis Gabor's (1900–1979) analytic signal theory, the Heyser spiral displays the concept of the complex plane in a single visual flash. Heyser was a scientist in the purest sense of the word, employed by NASA, and audio was his hobby. When I first met Richard C Heyser in the mid-1960s, Richard worked for Jet Propulsion Laboratory as a senior scientist. He invited me to his home to see his personal laboratory. The first thing he showed me on his Time Delay Spectrometry (TDS) equipment was the Nyquist plot of a crossover network he was examining.

I gave the display a quick look and said, "*That looks like a Nyquist plot!*"

He replied, "*It is*."

"*But*," I said, "*no one makes a Nyquist analyzer*."

"*That's right*," he replied.

At this point I entered the modern age of audio analysis. It was a revelation to watch Dick tuning in the signal delay between his microphone and the loudspeaker he was testing until the correct bandpass filter Nyquist display appeared on the screen. Seeing the epicycles caused by resonances in the loudspeaker, and the passage of non-minimum phase responses back through all quadrants, opened up a million questions.

Heyser's work led to loudspeakers with vastly improved spatial response, something totally unrecognized in the amplitude-only days. Arrays became predictable and coherent. Signal alignment entered the thought of system designers. Heyser's Envelope Time Curve (ETC) technology resulted in the chance to meaningfully study loudspeaker–room interactions.

Because the most widely taught mathematical tools proceed from impulse responses, Heyser's transform is perceived "through a glass darkly." It is left in the hands of practitioners to further research into the transient behavior of loudspeakers. The decades-long lag of academia will eventually apply the lessons of the Heyser transform to transducer signal delay and signal delay interactions.

I hold Harry Olson of RCA in high regard because, as the editor of the *Journal of the Audio Engineering Society* in 1969, he found Richard C. Heyser's original paper in the wastebasket; it had been rejected by means of that society's inadequate, at that time, peer review system.

2.9 Calculators and Computers

Richard C. Heyser gave us the instrumentation and Tom Osborne of Hewlett-Packard gave us the mathematical tools to begin to understand what we were actually doing in both electronics and acoustics as applied to sound reinforcement systems. Back in the 1960s, we utilized test equipment from Hewlett-Packard, General Radio,Tektronics, and Bruel and Kjaer. I purchased one of the very first Hewlett-Packard 9100 computer calculators designed by Tom Osborne. The 9100's transcendental functions, memory, and print out facilities led to lengthy acoustic design algorithms. Many of these same algorithms are still used in today's computers.

When the HP 35 handheld calculator, so named for its thirty-five keys, appeared we immediately put it to use in our teaching. The proliferation of easy-to-use, accurate software such as Mathcad, Matlab, and Mathematica in today's computers encourages engineers to explore more precise avenues of design. It was of interest to me that a correspondent, to the best Listserv in the business, stated to his peers that he had entered the audio industry via digital equipment and its operation, and was at a loss to understand analog audio.

I first explored the magic of communication via a crystal radio, then the vacuum tube technology, followed by transistors, integrated circuits, and now fully digital equipment electronically; it is comforting to turn to the acoustic side of sound reinforcement as an old familiar friend. When, as I expect in the not too distant future, the reinforcement system reaches the human brain sans passage via air the cycle will be complete.

2.10 The Meaning of Communication

The future of audio and acoustics stands on the Shoulders of the Giants that we have discussed, and numerous ones that we have inadvertently overlooked. The discoverers of new and better ways to generate, distribute, and control sound will be measured consciously or unconsciously by their predecessor's standards. Fad and fundamentals will be judged eventually, and put into their proper place. Age councils that "the ancients are stealing our inventions." Understanding an old idea that is new to you can be as thrilling as it was to the first person to make the discovery.

The history of audio and acoustics is the saga of the mathematical understanding of fundamental physical laws. Hearing and seeing are illusionary, restricted by the inadequacy of our physical senses.

That the human brain processes music and art in a different hemisphere from speech and mathematics suggests the difference between information, that can be mathematically defined, and communication that cannot. A message is the flawless transmission of the text. Drama, music, and great oratory cannot be flawlessly transmitted by known physical systems. For example, the spatial integrity of a great orchestra in a remarkable acoustic space is today, even with our astounding technological strides, only realizable by attending the live performance. The complexity of the auditory senses defies efforts to record or transmit it faithfully.

The devilish power that telecommunications has provided demagogues is frightening, but shared communication has revealed to a much larger audience the prosperity of certain ideas over others, and one can hope that the metaphysics behind progress will penetrate a majority of the minds out there.

That the audio industry's history has barely begun is evidenced every time one attends a live performance. We will, one day, look back on the neglect of the metaphysical element, perhaps after we have uncovered the parameters, at present easily heard but unmeasurable, by our present sciences. History awaits the ability to generate the sound field rather than a sound field. When a computer is finally offered to us that are capable of such generation, the question it must answer is, "how does it feel?"

2.11 Historical Notes

When I first ventured into audio in 1951 with nothing more than a background in amateur radio (Ham) and some exposure to engineering in general, at the university level, it was interest in classical music and its reproduction that led the way.

Fortunately, the postwar revolution in audio was at its height and our operation of a small but eclectic "hi-fi" shop, The Golden Ear, led to personally meeting many of the pioneers of that movement. Paul Klipsch, Rudy Bozak, Avery Fisher, Frank McIntosh, Herman Hosmer Scott, Saul Marantz and others, literally came to our little shop to help us sell the products they had designed and manufactured. Purdue University faculty and students were among our customers; listening to the conversations between these customers and the manufacturers participating soon made us aware of the names and fame of the communication industry's giants, as well as the early acoustic giants such as Rayleigh, Sabine, and Helmholtz.

In 1951, there was no such thing as the Internet and the only way to get to know these giants was to haunt old bookstores and gradually collect their written works. Our book collecting continued and gradually grew to over 750 volumes on audio and acoustics.

In late 1958, I joined Klipsch & Associates as vice president which gave me the chance to travel nationally and internationally, including sound demonstrations at the Brussels World's Fair, and at the American National Exhibition in Moscow, where we spent two-and-half months during the summer of 1959.

In late 1959, I joined Altec Lansing where I was privileged to work with three unique innovators in audio technology: John Hilliard, Art Davis, Jim Noble, and a host of employees that had originally been part of ERPI, the motion picture service division of Western Electric. These were men who had known Wente and Thuras, Harvey Fletcher, Black, Bode, Nyquist, and Jim Lansing prior to his tragic early death. I was a vice president of Altec when I left in December 1972 to start Synergetic Audio Concepts (Syn Aud Con).

The next twenty-three years were spent teaching classes in the basics of audio and acoustics, consulting and writing texts including *Sound System Engineering*, *Acoustic Tests and Measurements* and *How to Build Loudspeaker Enclosures*, plus many papers for the *Journal of the Audio Engineering Society*, *Audio Magazine*, and other popular publications, in addition to hundreds of articles in the Syn-Aud-Con Newsletters.

I have been led to these reminiscences because of the usefulness of knowing how we got here, as the only sure guide to where we might be going. Therefore, I hope you will not ignore our attempts to share these underlying concepts with you to encourage you to become a professional in the communication industry. Many of the best of tomorrow's innovations will be found by studying the best of the old, combined with the new materials and techniques of the present. "The ancients are still stealing our inventions" is all too true and many papers written at the turn of the Twentieth Century have ideas not able to be implemented when the article was written, but are now possible.

Chapter 3

Sound and Our Brain

by Don Davis

3.1 The Human Brain 19
3.2 The Current Era 19
3.3 Unexpected Validation 22
Bibliography 22

3.1 The Human Brain

Over the course of my lifetime my audio experiences ran from using a tickler on a crystal in order to hear audio from a local radio station, to the vacuum tube era, transistors, integrated circuits, to today's digital wizardry. Back in the early 1970s, it was clear that digital circuitry would prevail. It took forty years for that insight to come to full fruition. I truly believe that now mental control of such systems will come to full fruition in the next 10 years. There are already in existence neuron chips, axon chips and synapse chips from the scientists at IBM working to understand the brain.

In this chapter it is pointed out the interweaving of some of the most useful audio engineering tools such as the Fourier transform, Gabor wavelets, and quantum research to the holographic behavior of human consciousness. For those of you young enough and well-educated enough to want to pursue this insight, this chapter provides guidance for the scientist, the engineer, and the technician. It leaves one free to choose the level of your study while still providing future material that is challenging, hence this chapter about the thinkers behind this revolution.

As it is reinforced several times in this book, "*Sharing an idea is communication, understanding the idea is metaphysical*." Buckminster Fuller.

The earliest evidence we can find where man first recognized that the brain was associated with consciousness was Alkmaion of Kroton (500 BC) who discerned, based on anatomical evidence, that the brain was essential for perception. The Hippocratic School preserved his view regarding the mental primacy of the brain to vision. Nothing in early medical systems claimed any intellectual capacity for the brain. The Egyptians, normally so fastidious in their care for the afterlife, heedlessly discarded the brain in funerary practices. Most early civilizations thought of the heart as primary to understanding.

In the early sixteenth century George Berkeley, Bishop of Cloyne, published a book in 1710 entitled, "*Treatise concerning the principle of human knowledge*." In this work he clearly pointed out the unreality of the external world by demonstrating the illusions of the senses as witnesses in this case. Over 100 years later Ernst Mach (1838–1916) carried Berkeley's "Du Motu" into Mach's principle which then led Albert Einstein's thought to dwell on relativity.

Whether space is an independent entity of its own, a number of relations between material objects or a mere subjective notion superimposed on the world is a question of long tradition and probably rooted in our common impression of the distinct differences between objects and the empty space between them.

If you truly enjoy "mind-bending" reading Newton, Euler, Kant and Helmholtz's views of objects, space, and relativity are there for the taking. Mach insists that absolute motion and absolute space, i.e. motion and space in them selves, reside only in our minds and cannot be revealed by experience, hence they are meaningless, idle metaphysical concepts and must not be used in a scientific context. In quantum mechanics, mass and spin are both measures of inertia. Therefore, there are inertial affects proportional to Planck's constant, such as spin–rotation coupling, which is due to the inertia of intrinsic spin.

3.2 The Current Era

We are currently afloat in "field" theories of the human brain and its distinction from mind and Mind. Francis Crick (of DNA fame), E. Roy John (brain research laboratories NYU school of medicine), and Johnjoe McFadden (University of Surrey) have all made significant contributions to the electromagnetic fields in human beings, as well as providing support for wider sweeps of these fields into conjunction with the universe's field such as suggested by K. H. Pribram in "*Proposal for a quantum physical basis for selective learning*."

What has all this to do with audio? Fourier, Gabor, Shannon et al. have had an immeasurable effect on audio. Pribram writes,

> *Gabor wavelets are windowed Fourier transforms that convert complex spatial (and temporal) patterns into component waves whose amplitude at their intersections become reinforced or diminished.*

Fourier processes are the basis of holography. Holograms can correlate and store a huge amount of information and have the advantage that the inverse transform returns the results of correlation into the spectral and temporal patterns that provide guidance in navigating our universe.

Let me share with you some of the jewels and lumps present in the papers of these researchers with the hope that you, too, possess ideas that might be triggered by these suggestions. One materialistic view by Richard Amoroso written as a criticism of John Bell (of Cern fame),

He mistakenly thought that mind was immaterial; it has only seemed this way for the last 300 years because the material aspects of the normal nounenon of consciousness have been hidden behind the nonlocal Planck barriers. If this were not so, minds would not be safe from external influences and mental problems would be the norm rather than the exception and strong-willed individuals would be easily able to harm weaker psyches.

I am left with the feeling that Mr. Amoroso hasn't peered from his academic cloistered ivory tower to observe current human behavior and the total mental manipulation of large groups of psyche. Another material view of consciousness is that of Francis Crick in his paper, "A framework for consciousness," written with Christop Koch,

> *There has been a great selective advantage in reacting very rapidly, for both predatory and prey; for this reason the best is the enemy of the good.*

Col. Jeff Cooper taught many successful SWAT team members that proper shooting speed was measured by about 10% of your fast shots missing the center of mass of your target.

Crick further wrote,

> *More recently we have supported this suggestion that in addition to a slower, all—purpose conscious mode, the brain has many 'zombic modes,' which are characterized by rapid and somewhat stereotyped response.*

Driving a race car to its limits is not an intellectual but a highly programmed reflex.

Crick also wrote,

> *The conscious system may interfere somewhat with the concurrent zombic mode…it seems to be a great evolutionary advantage to have zombic modes that respond rapidly, in a stereotyped manner, together with a slightly slower system that allows time for thinking and planning more complex behavior.*

Col. Cooper's response to this was his famous color-coded awareness. His students trained to a zombic mode by over 10,000 repeats of the necessary physics of the quick reasonably accurate shot. The color code was designed to get them out of the unalert – unready "white" condition to a relaxed "yellow" where one looked about to see if all was normal. Where something wasn't normal then consciousness rose to a condition "orange." Orange alertness can go to "red" when danger is now imminent—the zombic mode is then released, but note, if truly alert, it's not a mindless act.

Driving a car in an unalert mode can be dangerous. If in yellow any darker patch on a winter road may lead to orange because it could be ice—you slow, it is ice—your zombic mode responds because you have raised your conscious awareness and are able to handle red.

It turns out that Mr. Crick is annoyed by physicists when he writes,

> *Many people have said that consciousness is "global" or has unity (whatever that is), but have provided few details about such unity. For many years Baars (MIT) has argued that consciousness must be widely distributed. We are not receptive to physicists trying to apply exotic physics to the brain, about which they seem to know very little, and even less about consciousness.*

It seems to this humble observer that since the physicists have measurably shown the presence of fields in the brain that employ the same components as fields in the universe that the possibility of similarity is not to be dismissed so cavalierly. I do suggest that Mr. Crick's paper is an exceptionally interesting read, especially where he sticks to his own expertise.

Turning to the other schools of thought such as John's, "The Neurophysics of Consciousness," global theories by Penrose and Hameroff invoking quantum mechanical concepts, and by Tononi and Edelman introducing measurements of system entropy and complexity have been proposed to account for the emergence of consciousness. Perhaps the most audacious and comprehensive approach to the explication of global intricate to process has been made by Pribram who contends that a holographic encoding of the nodes of interference patterns contains all of the information about the environment described by Gabor's elementary functions, quantum of information. He has explicated the relations between entropy, chaotic attractors, and the organization of the ensembles of Gabor's Quanta.

John says,

> *In critical observations about dependence of information encoding upon synchronicity within a region, dispersion of features extractors across brain regions, coherence among regions and the relevance of statistical considerations to explain brain functions cannot be reconciled with the hypotheses based upon discrete processes in dedicated cells. It is not plausible that a neuron can encompass the global information content of the multidimensional system to which it belongs.*

This paper is backed up by positron emission tomographs of living human brains and a complete schematic of the sources of some of the fields. Again a paper worth reading completely. The 216 listings in the bibliography are priceless to the researcher.

We now come to Johnjoe McFadden's "Synchronous firing and its influence on the brains' electromagnetic field."

> *It has been known for more than a century that the brain generates its own electromagnetic [EM] field. The electrical field at any point in the brain will be a superposition of the induced fields from all of the neurons in the vicinity (superimposed on the fields generated by ion movement) and will depend on their firing frequency, geometry and the dielectric properties of the tissue. Direct measurements of local field potentials within the human brain tissue has become possible in patients who for therapeutic reasons, have had EEG recordings obtained from sub dural cortical or in-depth electrodes implanted in their cerebral cortex. A striking feature of EEG is the differences in electrical activity from electrode to electrode even on less than 1 mm apart indicating that the brain generates a highly structured and dynamic extracellular electrical field. Further measurements revealed that the human (and animal) brains therefore contain a highly structured (in time and space) endogenous extracellular EM field, with a magnitude of up to several tens of volts per meter.*

The superposition principle states that for overlapping fields, the totally EM field strength at any point is an algebraic sum of the component fields acting at that point.

Like all wave phenomenon field modulations due to nerve firing will demonstrate constructive or destructive interference depending on the relative phase of the component fields. Temporal random nerve firing will generally generate incoherent field modulations leading to destructive interference and zero net fields. In contrast synchronous nerve firing will phase lock the field modulations to generate a coherent field of magnitude that is the vector sum (the geometric sum – taking into account the direction of the field) and its components.

Why such emphasis on this subject in a book on sound system engineering? It is my belief that nanotechnology and digital technology will within a generation be speaking directly to an audience's brain sans the acoustic link.

In humans, the strongest evidence for the sensitivity of the brain to relatively weak EM fields comes from the therapeutic use of transcranial magnetic stimulation (TMS). McFadden states:

> *TMS has been shown to generate a range of cognitive disturbances in subjects including: modification of reaction time, induction of Phosphenes, suppression of visual perception, speech arrest, disturbances of high movements, and mood changes.*

Repetitive TMS is subject to strict safety guidelines to prevent inducing seizures in normal subjects. It is striking how well shielded the human brain is from electrical signals cell phones et al, but such tremendous sensitivity to magnetic influences. The suggestion naturally arises as to what external fields can communicate with the human brain's internal fields directly.

McFadden's "Conscious electromagnetic information field (CEMI) field theory" states:

> *Digital information within neurons is pooled and integrated to form an electromagnetic field in the brain. Consciousness is the component of the brain's electromagnetic information field that is transmitted to motor neurons and is thereby capable of communicating its state to the outside world.*

McFadden further notes that:

> *Thanks to the Fourier transforms and wavelet transforms, linear superpositions or Laplacian information technology already exploits the advantage of EM information transmission in optical fiber communications.*

McFadden, unlike many of the more materialistic of his brethren, appreciates the work of physicist, Bernard J. Baars, and finds that his work and Baars support one another.

In this CEMI field theory, we are not simply automatons that happen to be aware of our actions, our awareness; (the global CEMI field) plays a causal role in determining our conscious actions.

Creating communication that results in conscious stimulation is our business. The B. Friedlaenders have also pointed out a similarity between their concept of relative inertia and induction effects in electromagnetism: just as a change in the magnitude of the current, or distance, will generate induction effects, only changes in velocity attractive will generate attractive, or repulsive, effects.

Strange as it is, quantum theory offers features which may be relevant to consciousness. One is that large collections of quantum particle/waves can merge into unitary coherent states of microscopic size and influence. Superconductors, Bose-Einstein condensates, and lasers are unitary states in which component atoms or molecules give up individual identity and behavior. Such coherent quantum states have been suggested to occur among brain proteins to provide unitary "binding" in vision and sense of self.

We now know that at very small scales, space and time are not smooth, but quantized. This granularity occurs at the incredibly small dimensions of the Planck scale (10^{-33}centimeters and 10^{-43} seconds). Penrose suggests thus everything is in reality particular arrangements of space-time geometry. Lee Smolin likens spin network volumes to Leibniz monads and suggests that self organizing processes at this level constitute a flow of time, raising the issue of whether the universe is in some sense alive.

3.3 Unexpected Validation

On December 20, 2011, IBM predicted mind reading machines. The announcement stated that IBM scientists are among those researching how to link your brain to your devices, such as a computer or a smart phone. IBM gave the examples of ringing someone up just by thinking it, or willing a cursor to move on a computer screen. They further state that biological makeup will become the key to personal identity, with retina scans, recognition of faces, or voices, used to confirm who people are rather than typing in passwords.

The metaphysics of such technical developments' ability to access our conscious thinking process (which this chapter on brain has discussed as electromagnetic energy) leaves open the possibility of two way transmission along the same paths. The current prediction that 80% of the population will have means of intercommunication with handheld devices such as I Pads and cellular telephones suggests that addressing large crowds will not require large sound systems but rather control of the individual devices carried by the majority of the population. Just as my use of digital voice recognition has allowed me to "type" these words (requiring an effort on my part to enunciate more clearly than was my normal pattern and allowing the program to learn the characteristics of my speech), so will we be required to obtain control of our often many layered consciousness when in the presence of mind reading devices.

We have had an exponential growth rate of hardware over recent decades, and it is now apparent that some of the software is beginning to grow exponentially as well. The earlier advent of neuron, axon, and synapse chips by IBM (see Chapter 5 *Digital Theory*) coupled to this most recent announcement regarding access to the internal signals of the human brain suggests that Orwell was only in error by 100 years in the means of human control to be accomplished in 2080.

Bibliography

B. J. Baars & K McGovern (1988), *A Theory of Consciousness,* New York: Cambridge University Press.

F. Crick & C. Koch, "A Framework for Consciousness", *Nature Neuroscience,* 6 (2), February 2003.

E. R. John, "The Neurophysics of Consciousness", *Brain Research Reviews* 39 (2002) pp. 1–28.

S. Martine-Conde, "A Review of Christof Koch's *The Quest for Consciousness*", *Psyche* 10 (2), September 2004.

http//www.merkle.com/humanMemory.html, No. 4, October 1988.

J. McFadden, "Synchronous Firing and Its Influence on the Brain's Electromagnetic Field", *Journal of Consciousness Studies*, 9 (4), 2002, pp. 23–50.

H. Moravec. "When will Computer Hardware Match the Human Brain?" *Journal of Evolution & Technology,* (1) (1998).

V. Vinge, "The Coming Technological Singularity: how to survive in the Post-Human Era", *Whole Earth Review*, Winter (1993).

Chapter 4

Psychoacoustics

by Don Davis

4.1 Motivations 25
4.2 Sound Reproduction 25
4.3 Is it Better to be Born Blind or Deaf ? 26
4.4 Recording Sound at the Eardrum 27
4.5 Psychoacoustics via a Metaphysical Foundation 27
4.6 Barks, Bands, Equivalent Rectangular Bandwidths (ERBs), Phons and Sones 28
    Phon Level 29
    Sones 29

4.1 Motivations

As I mentioned elsewhere in the book, my original motivation for pursuing audio studies came from extensive listening to classical music. This listening begun, in a serious way in 1949, and by 1952, I was engaged in the custom construction of what was then called high fidelity systems. Parallel to these activities I had the opportunity to hear many large orchestras and many fine artists due to the fact that we lived in a University town. In attempting to match the sound level, low distortion, and wide frequency range of a large orchestra we encountered many unmeasurable qualities that led us to understand the difference between "high fidelity" and fidelity. The difference was vast then, and it is still vast today.

Richard C. Heyser once remarked that the scene boards used on motion picture sets (the scene technical data is written on the board and a sound clap is made with a hinged section of the board) provided the necessary impulse data to let us, years later, if we ever acquired the ability, to reproduce the acoustic environment present on that soundstage.

Some of the better classical recordings, such as those made by EMI years ago, allow the listener to hear the room's reverberation and unique ambient noises. The spaces between crescendos can have important emotional content, especially in operas. There may be many significant signals in modern-day recordings that contain emotional content, if not processed out by modern digital technology.

The world's psychoacoustians have made great strides in aiding digital recording technology to remove meaning and emotion while retaining information. It is indeed true that Shannon ignored both meaning and emotion as irrelevant to his goal of transmitting and preserving information, but the technology has hidden within it possibilities as yet unimagined. None of these comments are intended to deprecate the remarkable gains that digital technology has provided us, but rather to encourage exploration into what constitutes true fidelity. I would encourage anyone making archival recordings today to never use less than 96 kHz sampling rate and twenty-four bit depth for two channel recordings

I am not a fan of "tested in the home," double-blind tests, etc. Over the decades, it has been shown that subtle flaws in recordings are detected by listening over and over again to favorite selections without any thought of conscious analysis. Often the realization will sneak in that the recording caused undue fatigue compared to other recordings of the same material. This has been true for both analog and digital recordings.

I have, in sixty years in audio, encountered a few individuals whose hearing, experience and intelligence provided exceptional judgmental capabilities that rank at the top of the Gaussian curve for human hearing. Such individuals can be an interesting guide to new listening experiences. A music critic detected in the recording of an older diva that a single high note had been dubbed in for her. He then identified the artist who had supplied the single note. A music critic who can do that demands our recognition and respect, both for his aural acuity and experience.

4.2 Sound Reproduction

Pundits, measurements, friends' advice, etc., can all serve as rough guides to seeking out a satisfactory reproduction system. The truly critical listener, however, is you. When listening to loudspeakers for the reproduction of music, previous experience plays an enormous role. True musical fidelity is in the live performance of gifted artists, and the greater the experience in hearing such artists the more likely the ability to detect unfaithful musical qualities in a loudspeaker.

The halls the artists perform in not only affect the acoustical quality of the signal they produce; but affect the sense the artist has of his own performance. Even partial reproduction of the performing environment in a recording is a rare occasion. Thus we begin the selection of a loudspeaker system for musical reproduction as a compromise.

I have read books, tested in the home reports, published laboratory reports, emotional evaluations; none of which ended up bearing any relation to what I heard when I had a chance to hear the loudspeakers they described. In my sixty years of involvement in audio systems I have found that the choices I made in my youth are not the choices I make in my maturity. In most cases this was not due to the technical advances made in the sixty years, but to my increasing experience in listening to loudspeakers. The all-horn loaded devices I so enjoyed in the 1950s, because of their dynamic range and directional control, was tempered by the reproduction of violins by the better direct radiators, coupled to suitable bass re-producers.

Surprisingly, the best electronics of sixty years ago are still the best electronics today, though not readily available as a consumer product. Interestingly, the professional monitor manufacturers produce the most reliable loudspeakers, which are least likely to fatigue with age, or be damaged by excessive levels. Horn loaded low-frequency loudspeakers lead by a substantial margin any other kind in the reproduction of transients, especially in piano music, percussion, and sound affects.

The best test of any loudspeaker system is the human voice. It's what we know best and is the easiest to judge for fidelity, simply by asking the listener to identify the talker. Remembering that the simple telephone circuit also allows this over a limited range, we need to add the tests of walking across the loudspeaker's pattern and listening to the tonal changes that occur with changing position. Signal delay and phase difficulties reveal themselves to the experienced listener with such changes in listening positions.

Pipe organ, coupled with orchestra recordings, reveal intermodulation distortion, when such distortion is present. In some early, very inefficient direct radiators, intermodulation removed the sound from higher frequency instruments, as the pipe organ pedals were utilized. This was not an electronic affect, but an electroacoustic effect.

In the final analysis when choosing a loudspeaker for critical listening purposes, select the one that sounds best to you, sans the advice of others. My best advice is to select the best one you can't afford because you will grow into it and growing into it is far better than growing past it.

The most reliable test of a loudspeaker is that you can listen to your favorite music by the hour. When your listening experience results in early fatigue and subconscious irritability, you have grown past your unit. In the purchase of an expensive loudspeaker system for your home arrange, with the dealer for its use in your home for several days prior to final purchase.

One of the joys of life in audio is when a highly experienced colleague is able to guide you to hear, over your loudspeaker, sounds previously inaudible to you that increase your listening pleasure.

Don't expect others to appreciate the choice you made. Engineers make radically different choices than musicians. Musicians upon hearing the harmonics will mentally replace the missing fundamental. Engineers will cringe at gross harmonic distortion whereas, musicians may not hear it. In a working environment such as a recording studio, the comparison of the monitoring loudspeaker and the live sound is but a step through the door to the studio, whereas in the home it's a trip across town to the concert hall.

Pass off as abused children anyone who would criticize your choice of a personal loudspeaker.

4.3 Is it Better to be Born Blind or Deaf ?

Helen Keller's statement on this subject:

> *I am just as deaf as I am blind. The problems of deafness are deeper and more complex, if not more important, than those of blindness. Deafness is a much worse misfortune. For it means the loss of the most vital stimulus—the sound of the voice that brings language, sets thought astir, and keeps us in the intellectual company of man.*

From an unknown author:

> *Let's focus on the blind/deaf question. Genius overcomes many difficulties. As evidence we have the pantheon of blind and deaf artists, ranging from Beethoven to Goya to Milton to Ray Charles. According to neurophysiologist and author, Oliver Sacks (in his book* Seeing Voices*), whether it's better to be blind or deaf depends on how old you are. For an adult, blindness and deafness are about equally problematic. But for a child, there is no question: it is better to be blind. Anyone who has had the opportunity to teach a deaf child knows this. Hearing is the primary channel through which we receive language, and all of those incoming words downloaded into our brains carry a wealth of emotional and cognitive apparatus that structures and empowers our imagination. Language is the mind's apposable thumb.*

Victor Peutz, who made a lifelong study of intelligibility and the percent of Articulation Loss of Consonants, often stated that it would be better to be born blind than deaf as language is necessary to the development of the intellect. The key word here is "born." Helen Keller was finally reached by Anne Sullivan when Helen made the connection from water out of a pump to the hand signals that Sullivan used for water. Helen, as an infant just before the illness that destroyed her visual and auditory senses, had experienced water.

For older adults it's a moot question. It's worth noting video images require only eight bits of digital information compared to audio which requires twenty-four bits for full fidelity. I have a neighbor whose hearing was restored via a cochlear implant. She was born with limited hearing. By the time she was 40 she was completely deaf, though she raised a family and operated a successful beauty shop by reading lips.

Modern hearing protectors are a very worthwhile investment especially when you're in the presence of impulsive sounds, such as gunfire, stapling guns, and various trip hammer devices. While extremely high-level non-impulsive sounds can cause discomfort and even temporary threshold shift in the human hearing, high-level impulsive sounds can cause loss on a single exposure.

4.4 Recording Sound at the Eardrum

Binaural recording has a history dating from the 1880s. True binaural recordings are two channels for reproduction through headphones. When Mead Killion, Eytomotic Research, developed wide-range probe microphones capable of being used at the human eardrum, I utilized them for measurement work in recording control rooms and for musical recordings in theaters with audience present. Using this new ITE (in-the-ear—at the eardrum) system, employing the pinna acoustic response recording techniques, we experienced for the first time in our lives, listeners unable to tell a live talker from a recorded talker when the recording is played back and the talker on the recording is also present in the listening room. The first time I listened to a playback made in our classroom I got up twice and went to the door to answer a knock, before realizing that the knock had come during the recording session. Harvey Fletcher in the SMPTE Journal Volume 61, September 1953 stated:

> *It is important to recognize the difference between a stereophonic system and a binaural system. The former system uses loudspeakers, but requires an infinite number of channels for perfect reproduction. The latter requires only two channels for perfect reproduction, but involves the use of a pair of head phones held tightly to the ears for each listener. All listeners with such a system can be given the illusion of sitting in the best seat in the concert hall.*

Most binaural recording today is done with an artificial or "dummy" head replicating the human head not only in average dimensions and details but also in approximate hardness and softness of skin and bone. Some of the recording heads also model the shoulders, and many have hair on the head, because all these details have an effect on the sound picked up by the two microphones.

These microphones are usually tiny omnidirectional condensers mounted at or near the entrance to the ear canals. Some designs place the microphones at the same location as the eardrums. In every case it is attempted to preserve the head related transfer functions (HRTFs), Fig. 4-2.

The notable difference between ITE recordings made in a live human head and more conventional techniques might be due to otoacoustic emissions or efferent stimuli that originates in the auditory cortex and terminates at the sensor of the organ of Corti.

Throughout its descending course the efferent auditory pathway interacts with the afferent auditory path through feedback loops and is identified with producing muscular and other reactions. The afferent pathway is the one from the cochlea to the brain areas associated with hearing. The efferent pathways are the ones desending from the brain back to the cochlea area. The descending (efferent) are known to have modulatory control of the medial efferent auditory system. (Nina Kraus at Northwestern University in Evanston Illinois) has written numerous publications addressing this subject.

The use of in-ear recordings allows later laboratory judgments about which direction undesired energy came from, and in the case of recording control rooms, shows a remarkable difference in envelope time curves from those made by conventional omnidirectional microphones. One obvious use for the better quality "dummy" heads is the ability to place them inconspicuously out in the audience so that a control room removed from the audience can have meaningful auditory input from the auditorium.

In the years when we made the original ITE recordings, the phase response between DAT recorders was not consistent. Today's solid-state digital recorders have the required phase stability to allow the interchange of recordings from one unit to the other via data cards.

One remarkable use of a precision artificial head was in the Acoustics Lab at Mercedes Benz in Stuttgart whereby they traced the noise sources and their directivity as part of a project to silence diesel engine noise in cars. I recently had the experience of riding in a Mercedes sedan where I could detect no difference in noise level from their gasoline fueled cars, inside and outside the car.

4.5 Psychoacoustics via a Metaphysical Foundation

I would like to take a moment of the reader's time to discuss psychoacoustics from a metaphysical, rather than a materialistic foundation. The measurements made to explore the vulnerability of humans to illusions, while valuable "leaves out all that is near and dear to our hearts."

Buckminster Fuller has said, *To share an idea is communication, to understand the idea is metaphysical.* As a young man I resisted musical training.

Military bands and marches were my only connection emotionally to music. At the university level, I encountered Vladimir Horowitz, large orchestras, and opera. The impact of Tosca sung by Maria Callis, the Mozart sonatas, and Beethoven's Ninth Symphony had lasting emotional effect on my state of mind.

At this period in my life, a close friend, a painter of unusually sensitive skill, pointed out that listening to music of a given period, while viewing paintings of the same time in history, and reading history of that period, led to a remarkably increased understanding of all three unobtainable when looking at only one of them.

Viewing the films made in the 1930s at Hitler's Nurenberg rallies, listening to the German National Anthem, and viewing the response of the audience, contrasts sharply with films made after the war of large crowds hearing the German National Anthem in Germany.

Artists, demagogues, political leaders, ministers, and teachers all communicate to their audiences, and on rare occasions, reach them metaphysically. The frenzies observed at rock concerts, the fury of Third World mobs aroused by a demagogue with a powerful sound system, and the connection of an inspired preacher to his audience through a large powerful sound system, are modern-day examples of psychoacoustics in the metaphysical sense.

From an engineering standpoint we know how to contribute to such events by properly enhancing the music, the speech, and the lighting. I recall one sound contractor being approached by an evangelical minister who requested he put more "soul" into the monitor. I didn't ask him to divulge this trade secret.

In Europe, "tone meisters" are taught technology, music and history. In historic venues the sound system engineer should emulate the medical physician by "first do no harm" as a goal before trying to introduce sound apparatus. The design goal should always be so unobtrusive that the audience does not realize a sound system is present until it is turned off unexpectedly and its presence is missed.

At the Opera house in Wiesbaden Germany, we congratulated the tone meister for the quality of the children's choir in Hansel and Greta without sound reinforcement. He replied "oh, but we did reinforce them," and then showed us an extremely clever loudspeaker system built into the arch of the stage opening, and used so sensitively as to be undetectable while reinforcing the children significantly.

We witnessed this particular performance while seated in the Kaiser's box, as guests of the tone meister. It was this experience that led me to remark that "I'd be willing to live in a class system but only if I could be King."

In today's world of tight budgets, compressed timescales, and isolation of the designer from the end-user, all of the above may be wishful thinking, but when possible, exceptional results can be obtained.

4.6 Barks, Bands, Equivalent Rectangular Bandwidths (ERBs), Phons and Sones

Psychoacoustics is employed in many areas of audio to manipulate the acoustic experience. Digital audio uses knowledge of phon levels and critical bandwidths to remove unnecessary data from being encoded into the digital stream. Filter design has utilized critical bandwidths as optimal numbers required to obtain desired equalization.

The concept of critical bandwidths first appeared, to the very best of my knowledge, in Harvey Fletcher's work with regard to masking. An article in "*Acustica*" in 1956 by H. Bauch entitled, "The Relevance of Critical Bands for the Loudness of Complex Sounds," was used by me in the 1960s to justify the bandwidth chosen in the design of equalizing filters.

The Bark scale ranges from 1 to 24 Barks, corresponding to the first twenty-four critical bands of hearing. The Bark scales are found in Table 4-1. Bark number, center frequency, and upper and lower −3 dB points are indicated. Critical bands shaped masking patterns should be seen as forming around specific stimuli in the ear, rather than being associated with a specific fixed filter bank in the ear. The equation for finding the bandwidth knowing the center frequency is:

$$BW = 25 + 75(1 + 1.4f^2)^{0.69} \tag{4-1}$$

where,

$$f = \frac{\text{Center frequency}}{1000}.$$

For example, using a center frequeny of 1080 Hz:

$$f = \frac{1080}{1000}$$

$$= 1.08$$

$$BW = 25 + 75(1 + 1.4(1.08)^2)^{0.69}$$

$$= 171\ \text{Hz}$$

To find the critical band z, use equation 4-2:

$$
\begin{aligned}
z &= \left(\frac{26.81 \times f}{1960 + f}\right) \qquad (4\text{-}2)\\
&= 9 \text{ Bark}
\end{aligned}
$$

Table 4-1. Bark Number, Center Frequency, and Upper and Lower −3 dB Points

Bark	f_c	f_l, f_u	Bark	f_c	f_l, f_u
0		0	13		2000
	50			2150	
1		100	14		2320
	150			2500	
2		200	15		2700
	250			2900	
3		300	16		3150
	350			3400	
4		400	17		3700
	450			4000	
5		510	18		4400
	570			4800	
6		630	19		5300
	700			5800	
7		770	20		6500
	840			7000	
8		920	21		7700
	1000			8500	
9		1080	22		9500
	1170			10500	
10		1270	23		12000
	1370			13500	
11		1480	24		15500
	1600				
12		1720			
	1850				

4.6.1 Phon Level

Fig. 4-1 shows the equal loudness contours from which phon levels are obtained. How widely individuals might vary from such curves, especially in the 2 to 4 kHz octave bands can be seen in Fig. 4-2 for pinnae responses. Fig. 4-3 compares one octave, one third of an octave, one sixth of an octave, ERBs and critical bands on a frequency versus bandwidth basis. ERBs are filters that are “Equivalent Rectangular Bandwidths” of critical bandwidth filters. To determine the ERB when the center frequency is known use:

$$
ERB = 0.108f + 24.7 \qquad (4\text{-}3)
$$

If we have a center frequency of 19,285 Hz, the bandwidth would be

$$
\begin{aligned}
ERB &= 0.108(19285) + 24.7\\
&= 2107.5 \text{ Hz}
\end{aligned}
$$

To find the center frequency of the band when the bandwidth is known use:

$$
f = \frac{(bw - 24.7)}{0.108} \qquad (4\text{-}4)
$$

For example, for a bandwidth of 142 Hz, the center frequency would be:

$$
\begin{aligned}
f &= \frac{(142 - 24.7)}{0.108}\\
&= 1028 \text{ Hz}
\end{aligned}
$$

Table 4-2 shows the equivalent rectangular bandwidth for the 41 center frequencies.

For a majority of uses, the one third octave-type filter (also the most available) will suffice. For overlapping critical bands the phon scales may be added on a power basis. Power basis is used instead of sound pressure level basis because of the definition of the phon as related to the sone.

4.6.2 Sones

We can take two adjacent one-third octave bands that lie on the 70 phon scale, add them as power ratios, and then convert the result to sones, Eqs. 4-3 and 4-4.

$$
\begin{aligned}
P_T &= 10\log\left(10^{\frac{L_{P1}}{10}} + 10^{\frac{L_{P2}}{10}}\right) \qquad (4\text{-}5)\\
&= x \text{ phons}
\end{aligned}
$$

where,
L_{P1} and L_{P2} are the individual sound levels in dB.

For example:

$$
\begin{aligned}
P_T &= 10\log\left(10^{\frac{70}{10}} + 10^{\frac{70}{10}}\right)\\
&= 73 \text{ phons}
\end{aligned}
$$

This allows a subjective sense of increased or decreased “loudness” on a scale thought comparable to human hearing. The chart shows that a change of 10 dB can be seen as doubling the sone value, hence twice the “loudness.” Going from one sone to two sones equals 10 dB by definition.

Table 4-2. Equivalent Rectangular Bandwidth Center Frequencies vs. Bandwidth

f_c	*bw*	f_c	*bw*
30	27.9	2284	271.4
61	31.3	2570	302.3
94	34.9	2889	336.7
131	38.8	3244	375.1
172	43.3	3640	417.8
218	48.2	4080	465.3
269	53.8	4571	518.4
326	59.9	5118	577.4
389	66.7	5727	643.2
459	74.3	6406	716.5
538	82.8	7161	798.1
625	92.2	8003	889.0
723	102.8	8941	990.3
831	114.4	9986	1103.2
952	127.5	11149	1228.8
1086	142.0	12446	1368.9
1236	158.2	13890	1524.8
1403	176.2	15498	1698.5
1589	196.3	17290	1892.0
1796	218.7	19285	2107.5
2027	243.6		

To determine the number of sones 73 phons is used:

$$S = 2^{\frac{P_T - 40}{10}} \tag{4-6}$$

or

$$S = 2^{\frac{73 - 40}{10}}$$

$$= 9.8 \text{ sones}$$

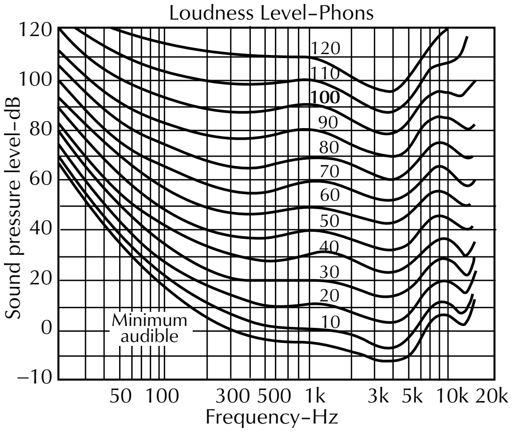


Figure 4-1. Equal loudness contours at various SPLs.

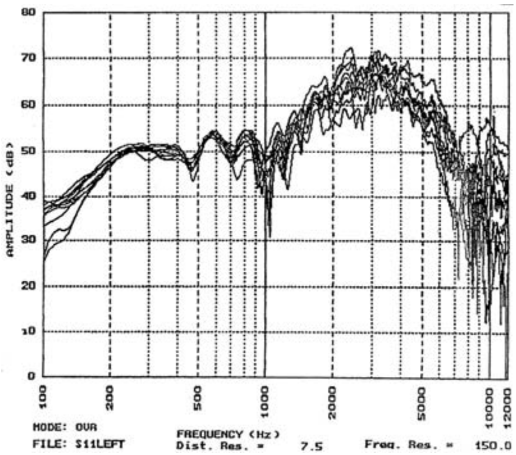


Figure 4-2. Pinnae responses.

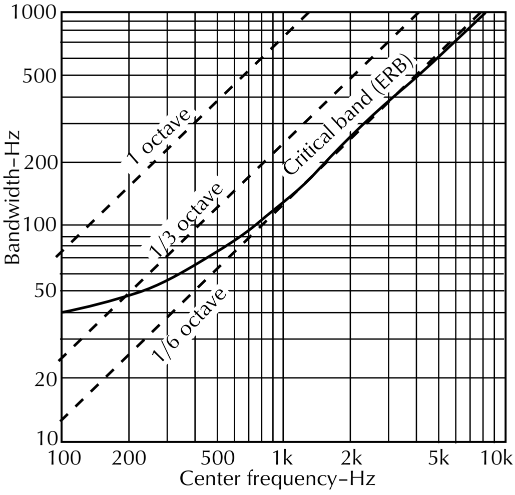


Figure 4-3. A plot of critical bandwidths of the human auditory system compared to constant percentage bandwidths of filter sets commonly used in acoustical measurements.

To go from sones to phons use:

$$P = 10\left(\frac{\ln(S)}{\ln(2)}\right) + 40 \qquad (4\text{-}7)$$

Therefore changing 8 sones to phons we find:

$$\begin{aligned} P &= 10\left(\frac{\ln(8)}{\ln(2)}\right) + 40 \\ &= 70 \text{ phons} \end{aligned}$$

Table 4-3 gives music levels vs. phons and sones.

Table 4-3. Relationship Between Music, Phons and Sones

Music	phons	sones
fff	100	64
ff	90	32
f	80	16
-	70	8
p	60	4
pp	50	2
ppp	40	1

Chapter 5

Digital Theory

by Don Davis

5.1 Shannon's Theory 35
5.2 Dynamic Range 36
  Cognitive Computing 37
  Digital Recording Techniques 37
  Frequency Dependent Case 39
  A Stochastic Process 40
5.3 The Steps from Art to Science 40
5.4 Moravec's Warning 41
5.5 Digital Nomenclature 42
5.6 What Is a Bit of Data? 43
  The Physical Dimensions of One Bit 44
  Reading Binary Numbers 45
  Text into Binary, Octal, Hexadecimal 46
5.7 Bayesian Theory 47
5.8 Planck System 47
5.9 Bits, Nats, and Bans 48
5.10 A Communication System 48
5.11 Holography 49
Bibliography 49

Theorists have postulated a holographic universe. Alex Harley Reeves wrote: "*The Bickenstein-Hawking conjecture asserts proportionality between surface area of a gravitationally closed structure like a black hole or a hypersphere, and its internal entropy.*" Richard C. Heyser commented during a lecture on Holography at the Audio Engineering Society meeting in Los Angeles that "*It pains me to give away in 10 minutes what it took me 10 years to learn.*" I remember thinking that it will take 10 years to assimilate everything he has shared in 10 minutes. See section 5.11 **Holography** for a discussion of Holography. The Claude Shannon Information Theory relates information to entropy. One Kelvin can be defined as a requirement of 1.38×10^{-23} J or 86.2 ueV of energy input per increase of the log state count by 1.0 Nat.

Combined these theories lead to the idea of a holographic universe in which each Planck unit of the surface of the universe carries one bit of information. These bits of information served to define what happens in all Planck volumes within the universe. Lord Rayleigh's, *The Theory of Sound*, had similar challenges to thought and the reliability of the evidence of the senses.

Kelvin says that ideal engines cannot exist, and Clausius says that ideal refrigerators can't exist.

Landauer's Principal states:

> *There is no machine whose sole effect is the erasure of information. There is a price to forgetting. One has to generate at least Kt (ln 2) to get rid of one bit of information.*

Shannon's great insight was that it is possible to associate entropy with any set of probabilities, Fig. 5-1.

5.1 Shannon's Theory

In our increasingly digital world we need to recall that the term bit "binary digit" first appeared in Claude Elwood Shannon's 1948 paper, "A Mathematical Theory of Communication," from the *Bell System Technical Journal*. (Given to Shannon by John W.Tukey, a Bell Lab coworker.)

The essential elements of Shannon's formula are:

1. Proportionality to bandwidth W.
2. Signal power S.
3. Noise power N.
4. A logarithmic function.

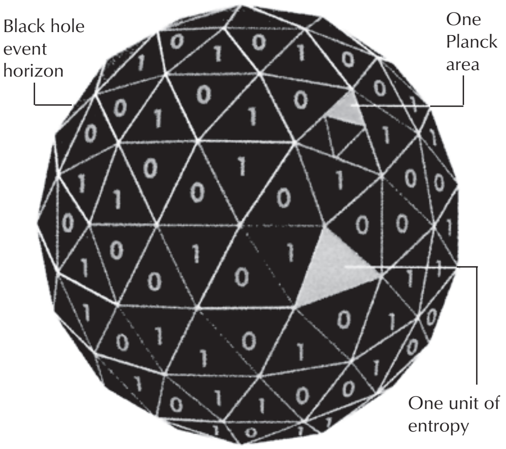


Figure 5-1. The information is stored as entropy on the area of the black hole's event horizon. Each bit (1s or 0s) corresponding to one unit of entropy, occupies four Planck areas. The area of the horizon is a measure of black hole's entropy.

$$C = W \times \log_2\left(1 + \frac{S}{N}\right) \qquad (5\text{-}1)$$

where,
C = bits of information that can be transmitted without error per second, Eq. 5-1.

Fig. 5-2 is a Mathcad printout of Information Theory.

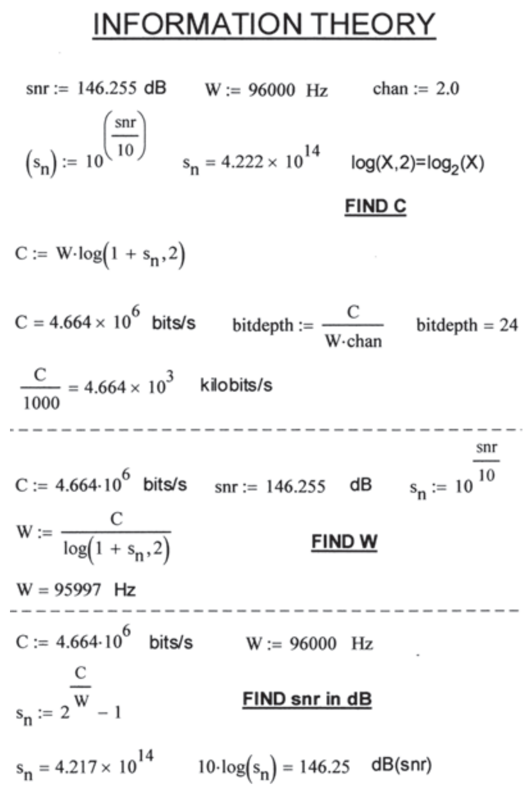
The seminal work of Alex Harley Reeves in 1937 when he patented PCM (pulse code modulation) quickly led to several conclusions.

Implicit in Reeves patent were two important principles:

1. An analog signal such as speech could be represented with accuracy by means of sufficiently frequent sampling and by quantizing each sample to one of a large number of predetermined levels, Fig. 5-3.
2. These samples can be transmitted with small probability of error provided the signal-to-noise ratio (SNR) is sufficiently large. This implied that such a channel could handle an infinite amount of information in an arbitrarily small bandwidth.

To encode an analog audio voltage into a digital audio number stream, analog-to-digital converters (ADCs) sample the waveform in both time—horizontal slices or samples—and in vertical voltage slices or quantization, usually expressed as some number of bits.

Professional audio devices typically use a 96 kHz sample rate, but other common rates are 32 kHz



INFORMATION THEORY

snr := 146.255 dB W := 96000 Hz chan := 2.0

$(s_n) := 10^{\left(\frac{snr}{10}\right)}$ $s_n = 4.222 \times 10^{14}$ $\log(X,2)=\log_2(X)$

FIND C

$C := W\cdot\log(1+s_n,2)$

$C = 4.664 \times 10^{6}$ bits/s $bitdepth := \frac{C}{W\cdot chan}$ $bitdepth = 24$

$\frac{C}{1000} = 4.664 \times 10^{3}$ kilobits/s

$C := 4.664\cdot10^{6}$ bits/s snr := 146.255 dB $s_n := 10^{\frac{snr}{10}}$

$W := \frac{C}{\log(1+s_n,2)}$ FIND W

W = 95997 Hz

$C := 4.664\cdot10^{6}$ bits/s W := 96000 Hz

$s_n := 2^{\frac{C}{W}} - 1$ FIND snr in dB

$s_n = 4.217 \times 10^{14}$ $10\cdot\log(s_n) = 146.25$ dB(snr)

Figure 5-2. Information Theory.

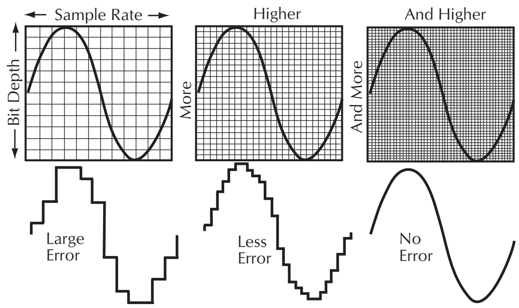


Figure 5-3. Effects of sampling rates on quality.

(broadcast), 44.1 kHz (audio CDs), and 48 kHz or 192 kHz for high-end recording or audiophile applications. Quantization, the number of bits used in professional equipment, is typically 24 bits with 16 bits for CDs.

"Shannon Space" for human hearing appears in Fig. 5-4.

5.2 Dynamic Range

Dynamic range theoretically for 24-bit quantization equals:

$$SNR = 10\log_{10}(6\times 2^{[(2\times bits)-2]}) \qquad (5\text{-}2)$$

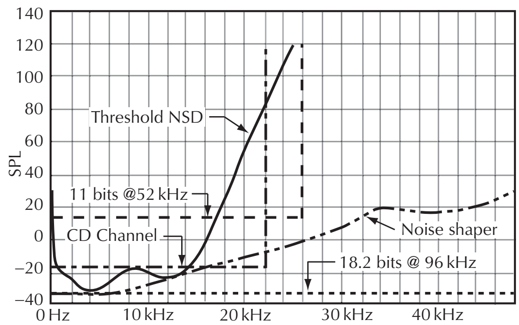


Figure 5-4. 'Shannon Space' for human hearing.

or 146.25 dB and for 16-bit quantization equals 98.09 dB.

To find the bit rate when the *SNR* is known use:

$$bits = \frac{\frac{\ln\left[\left(\frac{10^{\frac{SNR}{10}}}{6}\right)\right]}{\ln 2}+2}{2} \qquad (5\text{-}3)$$

Indeed digital recordings and electronics can in theory deliver increased dynamic range. Like distortion figures, dynamic ranges in real life sound reinforcement systems, while desirable, are usually constrained by the acoustic possibilities.

Consumer digital devices, cell phones, music players etc. are low-cost throw-away consumer items in today's market place. In professional sound systems, that operate in the presence of live audiences, battery operated analog backup systems for at least safety purposes, i.e. voice communication during a power failure, are still necessary adjuncts to a fully digital system. Latency problems also deserve special attention as many digital devices contain significant signal delay.

The digital world still falls short of what man can do inasmuch as the human brain has a storage capacity of from 10^{15} to 10^{17} bits of information and a processing rate of 100,000 Teraflops per second. The modern computing devices have reached the stage of being on a par with the brain of a Guppy, Fig. 5-5. The computer approach is a linear approach to a very nonlinear system called thinking. Humans have the very real ability to process very nonlinear information, distorted information, and even the ability to draw correct conclusions from false information.

In discussing neural networks David J. C. MacKay in his book, *Information Theory Inference and Learning Algorithms*, points out that digital devices suffer from:

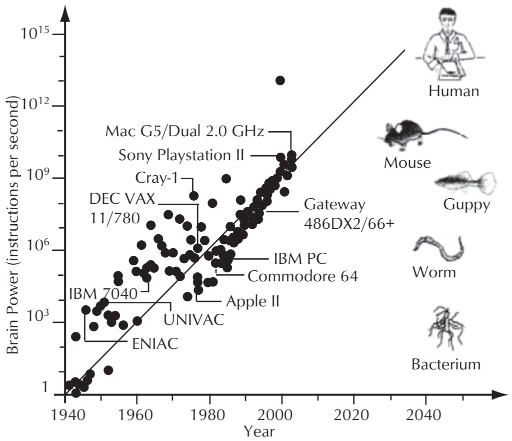


Figure 5-5. Raw Computing Muscle, as exemplified by a plot of 120 top machines of their time since 1940, is today on par with the brain of a guppy. It may reach human equivalent around 2040.

1. Address space memory is not associative.
2. Address space memory is not robust or fault tolerant.
3. Address-based memories are not distributed.

In the case of biological memory:

1. Biological memory is associative memory. Recall is content addressable.
2. Biological memory is error tolerant and robust. For example: "An American politician who was very intellectual and whose political father did not like broccoli" leads many people to think of President Bush (remember the author of this book is British) even though one of the cues contains an error.
3. Hardware faults can be tolerated. Memory often persists through brain damage.
4. Biological memory is parallel and distributed, and has a remarkable ability to work through loops.

The above is not to deprecate digital devices, but it merely intends to make the point that a live operator with multiple backup capabilities is wise insurance.

5.2.1 Cognitive Computing

August 18, 2011, IBM announced a series of chips that would allow a computer with processors that mimic the human brain's cognition, perception, and action abilities. It is described as:

> *The first cognitive computing core that combines computing in the form of neurons, memory in the form of synapses, and communications in the form of axons all working in silicon, and not PowerPoint. These chips can enable biological 'senses' such as sight, sound, smell, and touch, and drive multiple motor modes while consuming less than 20 W (of power) and occupying less volume than a 2L bottle of soda, and weighing less than 3 pounds.*

IBM hopes:

> *To weave the building blocks together into a scalable network and progressively scale it to a mammalian scale system with 10,000,000,000 neutrons, 100,000,000,000,000 synapses, all while consuming 1 kW of power and fitting in a shoebox sometime between now and the year 2018.*

Many brain researchers make a distinction between the brain and "Mind," and the role of each in consciousness. (See Chapter 3 *Sound and Our Brain* for further clarification of these distinctions).

5.2.2 Digital Recording Techniques

In 1928, Harry Nyquist wrote that sampling a signal at more than twice the desired bandwidth was a necessary limit on digital signaling. It is an error to say that it should be equal to twice the necessary bandwidth; it must be greater than the desired bandwidth.

Shannon termed the zeros and the ones as bits (binary digits), and employed the logarithmic base 2 in his calculations. From this sprang the concepts of sampling rate (samples per second), and quantization of the amplitude in bits. In audio recordings the sampling rate multiplied by the time in seconds, multiplied by the quantum value in bits, multiplied by the number of channels, divided by the channel reciprocal multiplied by bits equals the file size, Fig. 5-4. Eight bits equal one byte, four bits equals a nibble, and sixteen bits equals a word. File sizes are normally stated in bytes.

An example of decimal numbers as digital code is given in Table 5-2 where it can be seen that the decimal number is the addition of the exponents related to base 2; conversely you can by repeated division find a digital code from a decimal number.

Logarithms are used with many bases: the Napieran base e, the Briggsian base 10 (used for bans in code breaking during World War II), and base 2 for information. In the physical science world, one "Nat" has the physical dimensions of a square two Planck lengths on a side. The world-renowned physicist, John Wheeler, declared reality was "*It's from bits*."

Digital file size can be calculated with the following equation:

$$file\ size = \frac{(Sr \times ts \times bits \times chan)}{\frac{1}{chan} \times bits \times 10^6} \tag{5-4}$$

where,
Sr is the sampling rate,
ts is the time in seconds,
bits is the quantum value. 8 bits equal 1 byte,
chan is the number of channels.

An example is; if we have a sampling rate of 44,100 Hz, a time of 60 s, 16 bits, and 2 channels, we would have:

$$\begin{aligned} file\ size &= \frac{(44,100 \times 60 \times 16 \times 2)}{\frac{1}{2} \times 16 \times 10^6} \\ &= 10.584\ \text{mb} \end{aligned}$$

To find the file depth use:

$$file\ depth = 2^{bits} \tag{5-5}$$

or

$$\begin{aligned} file\ depth &= 2^{16} \\ &= 6.554 \times 10^4 \end{aligned}$$

To find the file depth in dB use;

$$file\ depth\ in\ dB = 20\log(file\ depth) \tag{5-6}$$

or

$$\begin{aligned} file\ depth\ in\ dB &= 20\log(6.554 \times 10^4) \\ &= 96.33\ \text{dB} \end{aligned} \tag{5-7}$$

The download time can be found by:

$$download\ time = \frac{\frac{file\ size}{modem\ speed \times ts}}{\frac{1}{chan} \times bits} \tag{5-8}$$

where,
ts is 1 for seconds, 60 for minutes, and 3600 for hours.

Assuming the modem speed is 0.056 mbps, the download time would be:

$$\begin{aligned} download\ time &= \frac{\frac{10.54 \times 10^6}{(0.056 \times 10^6) \times 60}}{\frac{1}{2} \times 16} \\ &= 25.2\ \text{min} \end{aligned}$$

Dr. Thomas Stockham made the very first 16-bit PCM recording in the United States in 1976 for the Santa Fe Opera on his Sound Stream recorder. When I first measured the "ringing" associated with the early digital recorders (antialiasing filters), Dr. Stockham was the only one I knew that had both understood and avoided this anomaly.

Studer, upon seeing the measurements, withheld their professional recorder until the problem was corrected; others failed to do so, which led, in our opinion, to some of the artifacts that so disturbed "sensitives" during that early period (1982). The Motion Picture Expert Group (MPEG) within the International Organization for Standardization (ISO), worked out a series of audio coding standards for storage and transmission of various digital media.

Digital versatile disc (DVD), High definition television HDTV, and ongoing competing methods make the description of each system chronologically challenging. Currently the Dolby AC-3 is the prevalent coding standard for the U. S. Phillips PASC (Precision Adaptive Subband Coding) is similar to the ISO/MPEG/1 layer 1; Sony has ATRAC (Adaptive TRansform Acoustic Coding) with its ability to manipulate psychoacoustic principles to both bit allocation and the time frequency mapping. Further digital details, mathematics and circuitry are in Chapter 22, *Signal Processing* and Chapter 25, *Putting It All Together.*

The broad parameters discussed here have not changed in the decades since Shannon's paper. Encoding and transmission techniques will continue to evolve, but the fundamental parameters continue to be usable guideposts when looking at devices and their claims.

All audio devices contain some latency. In digital devices such as crossovers (inadvertent delay) and deliberate (delay lines) can be acoustically significant in live sound reinforcement. In one sound system I was hired to evaluate, the delay in a digital crossover was 30 ms and rendered signal alignment a matter of putting one part of the loudspeaker system physically 30 ft behind the other part. Numerous occasions occurred wherein measuring the time domain behavior of a system left us with a blank screen on the analyzer until we remembered the fundamental rule to always measure globally before measuring specifically.

The sound system engineer must keep in mind that we start with an acoustic signal in an acoustic environment and end up with an acoustic signal in an acoustic environment in a vast majority of cases. In broadcasting and recording up to 90% of the actual signal material can be removed in the processing of the digital signal at hand. These were some of the early lessons in using digital equipment. Antialiasing filters, i.e. brick wall filters with attenuation rates as high as −146 dB per octave, were tried by some early CD manufacturers, resulting in a group delay at 20 kHz of 1 ms, and a relative phase shift of some 3000° which was clearly audible. The problem was finally solved and then only on replay by the use of over sampling techniques that moved the aliasing frequency from 22 kHz up to as high as some cases as 154 kHz. This allowed controlled filter design well away from the desired pass band.

Richard C. Heyser liked to point out how much the human listener can detect that can't be measured. He said:

> *The end product of audio is the listening experience. The end product is a result of perception and cognition and evaluation processes occurring in the mind. What do we know about such processes—the answer is very little.*

Heyser went on to elaborate:

> *My own research into nonlinear behavior has caused me to introduce three divisions to what is universally spoken of as perception. These are the divisions of:*
>
> 1. *Sensory contact and stimulus.*
> 2. *Association of stimulus with memory and past experience and ongoing stimuli of other nature.*
> 3. *Evaluation of stimulus in light of ongoing experience I call them perception, cognition, and valuation. We can like something today and not like the same tomorrow even though the program and the stimulus are essentially identical in both cases. The perception was unchanged, but cognition and evaluation were altered.*

Pre-emphasis

The signal-to-noise ratio can be improved by using high-frequency pre-emphasis. The choice of the pre-emphasis characteristic must be made with care. The curve for pre-emphasis is designed making assumptions about the program content spectrum.

Dither

At low levels, only a small number of states are available. This can lead to audible distortion such as the decay of a piano note. It has been found that adding a small amount of random noise significantly improves the perceived quality. This noise "dithers" the LSB and can be regarded as the digital equivalent of bias in magnetic recording.

Aliasing

It is important that the sampling process is protected from out-of-band frequencies. It was in the design of antialiasing filters in the early days of digital audio that artifacts became audible to the listener.

Over-sampling

A common technique to reduce the burden on the filters is over sampling. By reading the data two, four, or eight times, the spurious frequencies are raised one, two, or four octaves; therefore, need less severe filters to attenuate them to insignificant proportions.

Bit rate reduction

Bit reduction allows more audio to be processed in a given time. Halving the audio bit rate on the hard disk system can double the number of audio signals that can be simultaneously output from it. This is done for economic reasons.

5.2.3 Frequency Dependent Case

Where the additive noise is not white or that the signal-to-noise is not constant with frequency over the bandwidth, the following equation can be used by treating a series of channels as narrow, independent Gaussian channels in parallel.

$$C = \left(\int_0^W \log_2\right)\left(1 + \frac{S(f)}{N(f)}\right) df \tag{5-9}$$

where,
C is channel capacity in bits per second,

W is the bandwidth of the channel in Hz,
$S(f)$ is the signal power spectrum,
$N(f)$ is the noise power spectrum,
f is the frequency in Hz.

These equations can be used to demonstrate how spread spectrum communication systems make it possible to transmit signals which are actually much weaker than the background noise level.

There are three predominant types of noise;

1. White—equal energy per hertz.
2. Pink—equal energy per octave.
3. Black—silence.

Errors in digital transmission are considered as noise, as well as the naturally occurring thermal noise, and the effects of radio frequency interference (RFI). Relations between bandwidth and time similar to the one found by Nyquist was discovered simultaneously by Karl Kupfmuller in Germany. And even more stringent analysis of the relation was carried out in Gabor's, *Theory of Communication*.

As was pointed out by Tuller (1949), a fundamental deficiency of the theories of Nyquist, Hartley, Kupfmuller and Gabor, was that their formulae did not include noise. The role of noise is that it sets a fundamental limit to the number of levels that can be reliably distinguished by a receiver. Each of these scientist-engineers had worked with the fundamentals of noise and had fundamental papers that included the earliest measurements and mathematical derivations, relative to noise, but chose to treat the analysis of a communications channel as noiseless. Claude Shannon's genius was to unite all these disparate theories into one.

5.2.4 A Stochastic Process

A system which produces a sequence of symbols (letters of the alphabet or musical notes) according to certain probabilities is called a stochastic process, and the special case of a stochastic process in which the probabilities depend on the previous events, is called a Markov process or a Markov chain. Of the Markov processes which might conceivably generate messages, there is a special class which is of primary importance for communication theory, these being what are called ergodic processes.

An ergodic process is one which produces a sequence of symbols which would be a poll taker's dream, because any reasonably large sample tends to be representative of the sequence as a whole. A truly reverberant auditorium is an excellent example of an ergodic process, in as much as, any one point of measurement would give the same result as any other point of measurement.

The world is full of non-Markov processes. A Markov process is one in which the future depends only on the conditions in the past. If your awareness of something is changed irrevocably by the introduction of some new piece of knowledge, then the altered awareness is non-Markovian.

In discussing the signal delay between a low-frequency driver and a high-frequency driver with a world-renowned psychoacoustic authority who had stated that 3 ms was inaudible, I asked had he walked the polar pattern. The lesson was non-Markovian as the polar response had been altered by the seemingly innocuous signal delay, while the amplitude etc. had not. Audio, acoustic, and digital measurement systems are helpless whenever nonlinear phenomenon is present. All our measurement systems are Markovian inasmuch as we expect the input to predict the output. All our measurement systems are dependent upon linear equations. This is not to imply that they are useless because the skilled and experienced operator can often read ambiguous data.

As a guide to a new listening experience, Dr. John Diamond's detection of serious flaws in the early CDs did lead to correction of the antialiasing problems, but not, unfortunately, to his being properly acknowledged as having brought it to the recording world's attention other than as a non-Markovian moment for many overconfident engineers.

5.3 The Steps from Art to Science

Art usually precedes science, i.e. the musical scale followed by the discovery of logarithms. Many of the present-day entertainment systems—disco, rock concerts, etc, satisfy listeners emotionally, and when they do so it is wise to search amid the cacophony for the psychoacoustic clues, if only for tolerance levels of audiences. If this sounds facetious, remember the FBI used rock music at high levels to attempt to disorient the victims at Waco, and L Rad systems have been employed to repel pirates. The delight with which engineers embrace digital audio is remarkable in as much as the foreseeable future dictates computers more than likely making them obsolete.

Prior to Shannon's Theory, bandwidth, signal power, and modulation types were well-established, but the word "bit" was seen in print for the first time in his paper. From 1948 to the present time the hyperbolic growth of digital communication has literally exploded to the point where it is feared, politicized, and omnipresent. Our TV, internet, computers, cell phones, motion pictures, and yes, even audio are the outcome of Shannon's Information Theory.

Thomas K. Landauer worked at Bell Communications Research and used the concept of viewing human memory as a novel telephone line that carried information from the past to the future. He came to the conclusion by actual measurements, that human beings remember very nearly 2 bits per second under all experimental conditions—visual, verbal, musical—two bits per second. Landauer's numbers are unusual due to their small size. Von Neumann's early estimate had been many orders of magnitude higher because of the brain's 10^{15} capacity. Landauer's estimate taken literally would suggest lifetime memory storage of 10^9 bits.

There are other ways to look at the brain. There are roughly 10^{50} synapses operating at about 10^{10} pulses per second giving an estimated 10^{16} synapses operations per second. The total energy consumption of the brain is about 25 W. The computation thinking part uses about 10 W—the remainder controls the pumps, etc. The MIPW of the brain has to give AI researchers pause. In the attempt to digitally automate mankind, the theoretical limits are fundamentals, like the speed of light and Planck's constant. In the meantime mankind's demagogues, aided substantially by TV and sound systems may bring civilization as we know it to an end. Hitler's giant rallies, Third World mob scenes, all required powerful sound systems as do more subtle corrupters of freedom and liberty. As you study these very simple equations and their vocabulary perhaps you'll be among the numbers that see the concept that will supersede them if proper progress is to continue. In any case, the simplicity that underlies the hardware of digital brings forth the complexity of future software to control it.

Information is ultimately constrained by the fundamental laws of physics. It should not therefore be surprising that physics and information share a rich interface. Richard Feynman's lecture in 1959, where he discusses storing and manipulating information on the atomic level, has now been brought to pass. In 2006, IBM announced a circuitry on a 30 nm scale which indeed makes it possible to write the *Encyclopedia Britannica* on the head of a pin. Ironically Feynman's speculative remark in 1959 is now just a marker of the current scale of computation. To make it clear how close this is to the atomic scale, a square with sides of length 30 nm contains about 1000 atoms.

5.4 Moravec's Warning

Moravec has warned us of the consequences of the digital age:

> *Advancing computer performance is like water slowly flooding the landscape. Half-century ago it began to drown the low lands, driving out human calculators and record clerks, but leaving most of us dry. Now the flood has reached the foothills, and our outposts there are contemplating retreat. We feel safe on our peaks, but at the present rate, those too will be submerged within another half-century. He proposed that we build an ark. For now, though, we must rely on our representatives in the low lands to tell us what the water is really like.*

Summarizing

Information is always conserved. If this were not true, it would violate energy conservation and the temperature of the universe would've risen to 10^{31} degrees in a fraction of a second. Physicists have found that "black holes" have event horizons. Simply put, the black hole horizon, by some mechanism, stores all the in-falling information. The entropy of a black hole, its area, and the in-falling information are correlated, Fig. 5-1.

There is thermodynamic entropy. It measures the number of different microscopic configurations that a system can be in. Shannon entropy is a measure of the information contained in a message. Thermodynamic entropy is measured in units of energy divided by temperature, whereas Shannon entropy is measured in bits. There exists a correlation between entropy and information–entropy is hidden information i.e. it is the Information in the hidden details of the system. The difference between information storage and information communication is only a difference in one's inertial frame of reference. Communication from point A to point B is ultimately just bit transportation i.e. (a form of storage) but in a state of relative motion. Likewise storage is just communication across zero distance (but through time).

"Coding High Quality Digital Audio" by J. Robert Stuart relates Shannon's formula to digital coding practices, Fig. 5-4. In another paper; "The Physics of Information" by F. Alexander Bais and J. Doyne Farmer, we find:

> *The theory of thermodynamics taken by itself does not connect entropy with information. This only comes about when the results are interpreted in terms of a microscopic theory, in which case temperature can be interpreted as being related to uncertainty and incoherence in the position of particles.*

The role of quantum theory in the relationship of information and entropy is properly explored in this paper.

What originally triggered my renewed interest in "Information Theory" back in 2008 was reading Prof. Chris Bissell's writings on Karl Kupfmuller:

> *If today, we recognize information along with energy and matter as a third fundamental building block of the world.*

I was left with the challenge to my thinking to bring information into a meaningful relationship with physics. This is being done today using Planck values and quantum mechanics.

The physical dimensions of one bit are:

1. One Nat of information has the area of a square exactly two Planck lengths on a side.
2. One Nat equals 0.693 bits.
3. One Planck length equals

$$L_P = \sqrt{\frac{\bar{h} \times G}{C}} = 1.616252(81) \times 10^{-35}\ \text{m} \tag{5-10}$$

where,
G is the gravitational constant,
$\bar{h}$ is the reduced Planck constant, $\bar{h} = \left(\frac{h}{2\pi}\right)$
C is the speed of light.

4. Planck area equals:

$$L_P^2 = \frac{\bar{h} \times G}{C} = 2.61223 \times 10^{-70}\text{m}^2 \tag{5-11}$$

$$1.0\ \text{Nat} = (2 \times L_P)^2 = 1.044909(2.61223) \times 10^{-69}\text{m}^2 \tag{5-12}$$

$$1.0\ \text{bit} = \frac{1.0\ \text{Nat}}{\ln 2} = 1.507485(2.61223) \times 10^{-69}\text{m}^2 \tag{5-13}$$

5. $\sqrt{1\ \text{bit}} = 3.88263 \times 10^{-35}\text{m}$ is the length of one side of the bit square.
6. $\sqrt{1\ \text{Nat}} = 3.232506(3.88263) \times 10^{-35}\text{m}$ is the length of one side of the Nat square.
7. $$\frac{bit_{P_l}}{Nat_{P_l}} = 1.201122 \times 2P_l = 2.402244\ P_l$$
8. $$\frac{(2.4P_l)^2}{2P_l^2} = 0.693\ \text{bits per Nat.}$$

The physical dimensions of one bit provide an insight into the use of the bit in physics. That such dimensions seem unimaginable today reminds us that less than a decade ago nanotechnology provided the same challenges. Frank Wilczek of MIT has written:

> *The wave patterns that describe protons, neutrons, and their relatives resemble the vibration patterns of musical instruments. In fact the mathematical equations that govern these superficially very different realms are quite similar.*

Some physicists today have described reality as "*It's from bits*." Where Newtonians saw the universe as some form of giant mechanism, and early computer enthusiasts saw the universe as a form of giant computer, it is increasingly viewed as some form of Information. The singularity problem provides interesting reading where those who are sharing deep thinking about the role of information in our lives describe the undesirable possibilities of artificial intelligences governing man.

5.5 Digital Nomenclature

Table 5-1 is a list of the most popular nomenclatures used in digital circuitry

Table 5-1. Digital Nomenclature

Bit	binary digit
Nibble	4 bits
Byte	8 bits
Word	16 bits
Double word	32 bits
Quad word	64 bits
Binary	base 2, $\log_2$
Octal	base 8, $\log_8$
Denary	base 10, $\log_{10}$

Table 5-1. (cont.) Digital Nomenclature

Hexadecimal	base 16, $\log_{16}$
MSB	most significant bit
LSB	least significant bit
Bit rate	bits × sampling frequency × channels
SNR	signal-to-noise ratio, $10 \times \log_{10} (6 \times (2^{(b - 2)})$
FLOP	floating point operations per second
MIPS	million instructions per second
MIPW	million instructions per watt
RISC	reduced instruction set computer
WAV	wave format
BWF	broadcast wave format
CPU	central processing unit
GPU	graphics processing unit
PCM	pulse code modulation
PDF	portable document format
ASCII	American Standard Code for Information Interchanges
CD	compact disc
DVD	digital versatile disc
MPEG	motion picture experts group
ISO	International Organization for Standardization
AAC	advanced audio coding
MDCT	modified discrete cosine transform
TNS	temporal noise shaping
HDTV	high-definition television
VBR	variable bit rate
CBR	constant bit rate
Codec	compressor/decompressor
IO	input – output
Fs	sampling frequency
$V/2^{bits}$	Quantization level
DSP	digital signal processor
ADC	analog to digital converter
DAC	digital to analog converter
ERB	equivalent rectangular bandwidth
HTML	hypertext markup language
BAUD	a rate defined as Xbaud × Y bits/baud = Z bits/s. (Baud is changes of state/s) Named after French engineer Jean Maurice Emile Baudot.

5.6 What Is a Bit of Data?

> *The transmission of communications signals is accomplished by means of a transmission of energy, generally of electromagnetic or of acoustic energy. In contrast to the case of power transmission, it is not energy itself which is of interest, but rather the changes in this energy in the course of time. The more complicated the function which represents, as a function of time, the change in voltage, current, pressure, or any other carrier, the greater is the amount of information carried by the transmitted energy.*(J. Ville)
>
> *Communications theory has up to now developed mainly on mathematical lines, taking for granted the physical significance of the quantities which figure in this formalism. But communication is the transmission of physical effects from one system to another, hence communication theory should be considered as a branch of physics. Thus it is necessary to embody in its foundations such fundamental physical data as a quantum of action, and the discreteness of electrical charges. This is not only of theoretical interest. With the progress of electrical communications toward higher and higher frequencies we are approaching a region in which quantum effects become all-important. Nor must one forget that vision, one of the most important paths of communication, is based essentially on quantum effects.* (D. Gabor*)*

Reading from the book *The Mathematical Theory of Communication* by Claude E Shannon:

> *If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one messages chosen is from the set, all choices being equally likely. As was pointed out by Hartley the most natural choice is the logarithmic function. Although this definition must be generalized considerably when we consider the influence of the statistics of the message and when we have a continuous range of messages, we will in all cases use an essentially logarithmic measure.*

The logarithmic measure is more convenient for various reasons:

1. It is practically more useful. Parameters of engineering importance such as time, bandwidth, number of relays, etc., tend to vary linearly with the logarithm of the number of possibilities. For example, adding one relay to a group doubles the number of possible states of the relays. It adds 1 to the base 2 logarithm of this number. Doubling

the time roughly squares the number of possible messages, or doubles the logarithm, etc.

2. It is nearer to our intuitive feeling as to the proper measure. This is closely related to (1) above, since we intuitively measure entities by a linear comparison with the common standards. One feels, for example, that two punch cards should have twice the capacity of one for information storage, and two identical channels twice the capacity of one for transmitting information.
3. It is mathematically more suitable. Many of the limiting operations are simple in terms of the logarithm but would require clumsy restatement in terms of the number of possibilities.

The choice of a logarithmic base corresponds to the choice of a unit for measuring information. If the base 2 is used, the resulting units may be called binary digits, or more briefly bits, a word suggested by J. W. Tukey. A device with two stable positions, such as a relay or a flip-flop circuit can store one bit of information. N such devices can store N bits, since the total number of possible states is 2^N and $\log_2 2^N = N$. If the base 10 is used the units may be called decimal digits. Fig. 5-6 gives the definition of binary.

Since $M = \log_{10}N/\log_{10}2 = 3.32\log_{10}N$, a decimal digit is about 3⅓bits. A digital wheel on a desk computing machine has ten stable positions and therefore has a storage capacity of one decimal digit. In analytical work where integration and differentiation are involved, the base e is sometimes useful. The resulting units of information will be called natural units (Nats). Change from the base a to base b merely requires multiplication by $\log_b a$.

Mathematically the number of bits is defined by:

$$\log_2 P = Bits \qquad (5\text{-}14)$$

and

$$2^{Bits} = P \qquad (5\text{-}15)$$

where,
P is the number of possibilities.

5.6.1 The Physical Dimensions of One Bit

One Nat of information has the area of a square exactly two Planck lengths on a side.
One Nat equals 0.693 bits,
One Planck length equals

$$L_P = \sqrt{\frac{\bar{h}\times G}{C}} = 1.616252(81)\times 10^{-35}\ \text{m} \qquad (5\text{-}16)$$

A. Mathematically:
The simplest numbering scheme possible, there are only two symbols:
1 and **0**

B. Logically:
A system of thought in which there are only two states:
True and **False**

C. Binary information is not subject to misinterpretation:

Black	**White**
In	**Out**
Guilty	**Innocent**

D. Variables or non-binary terms:

Somewhat	**Undecided**
Probably	**Not proven**
Grey	**Under par**

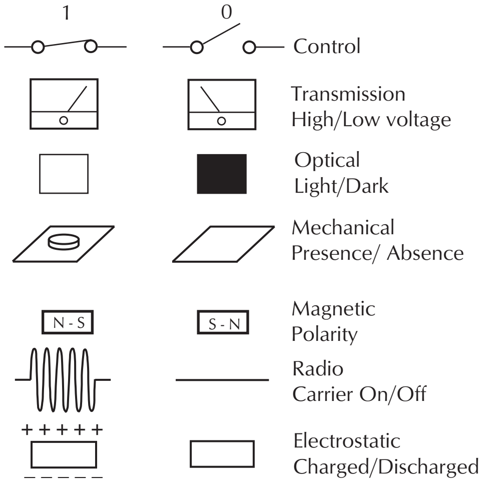


Figure 5-6. Binary choices.

where,

G is the gravitational constant,

$\bar{h}$ is the reduced Planck constant, where $\bar{h} = \left(\frac{h}{2\pi}\right)$

C is the speed of light.

Planck area equals

$$L_P^2 = \frac{\bar{h}\times G}{C} = 2.61223\times 10^{-70}\,\text{m}^2 \qquad (5\text{-}17)$$

$$1.0\ \text{Nat} = (2\times L_P)^2 = 1.044909(2.61223)\times 10^{-69}\,\text{m}^2 \qquad (5\text{-}18)$$

$$1.0\ \text{bit} = \frac{1.0\ \text{Nat}}{\ln 2} = 1.507485(2.61223)\times 10^{-69}\,\text{m}^2 \qquad (5\text{-}19)$$

$\sqrt{1 \text{ bit}} = 3.88263 \times 10^{-35}\text{m}$ the length of one side of the bit square

$$\sqrt{1 \text{ Nat}} = 3.232506(3.88263) \times 10^{-35}\text{m} \text{ the length of one side of the Nat square} \tag{5-20}$$

$$\frac{bit_{P_l}}{Nat_{P_l}} = 1.201122 \times 2P_l = 2.402244 \text{ P}_1 \tag{5-21}$$

$$\frac{(2.4P_l)^2}{2P_l^2} = 0.693 \text{ bits per Nat.} \tag{5-22}$$

5.6.2 Reading Binary Numbers

Table 5-2, Binary to Decimal to Hexadecimal to Octal allows you to see the essential simplicity of binary coding. Pick any decimal number and see how the 1s add up on the exponential scales at the top of the chart. For example the decimal number 27 in binary is found to be 16+8+2+1 =27. The hexadecimal number 1B is in the second rotation through the hexadecimal encoding, and octal number 33 is in the third rotation for octal encoding, see Table 5-2. Fortunately today many scientific calculators include easy conversions from decimal to octal to hexadecimal to binary code. Children can bring this kind of coding home from school for your help in solving their problems in high school mathematics classes. Cryptography is replete with many system bases including the Ban that was used in World War II by the English code breakers, see Fig. 5-7 defining Bits, Nats, and Bans. Digital test equipment is expensive can be complex and requires training in its use. We increasingly see requirements in specifications demanding certification of the engineers setting these systems into operation.

Table 5-2. Binary to Decimal to Hexadecimal to Octal

		Binary						
MSB					**LSB**			
2^5 (32)	2^4 (16)	2^3 (8)	2^2 (4)	2^1 (2)	2^0 (1)	**Decimal**	**Hex**	**Octal**
0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	1	1	1
0	0	0	0	1	0	2	2	2
0	0	0	0	1	1	3	3	3
0	0	0	1	0	0	4	4	4
0	0	0	1	0	1	5	5	5
0	0	0	1	1	0	6	6	6
0	0	0	1	1	1	7	7	7
0	0	1	0	0	0	8	8	10
0	0	1	0	0	1	9	9	11
0	0	1	0	1	0	10	A	12
0	0	1	0	1	1	11	B	13
0	0	1	1	0	0	12	C	14
0	0	1	1	0	1	13	D	15
0	0	1	1	1	0	14	E	16
0	0	1	1	1	1	15	F	17
0	1	0	0	0	0	16	10	20
0	1	0	0	0	1	17	11	21
0	1	0	0	1	0	18	12	22
0	1	0	0	1	1	19	13	23
0	1	0	1	0	0	20	14	24
0	1	0	1	0	1	21	15	25
0	1	0	1	1	0	22	16	26
0	1	0	1	1	1	23	17	27
0	1	1	0	0	0	24	18	30

Table 5-2. (cont.) Binary to Decimal to Hexadecimal to Octal

	0	1	1	0	0	1	25	19	31
	0	1	1	0	1	0	26	1A	32
	0	1	1	0	1	1	27	1B	33
	0	1	1	1	0	0	28	1C	34
	0	1	1	1	0	1	29	1D	35
	0	1	1	1	1	0	30	1E	36
	0	1	1	1	1	1	31	1F	37
0	1	0	0	0	0	0	32	20	40

MSB most significant bit LSB least significant bit

5.6.3 Text into Binary, Octal, Hexadecimal

ASCII printable characters

Codes 32–127 are common for all the different variations of the ASCII table; they are called printable characters and represent letters, digits, punctuation marks, and a few miscellaneous symbols. ASCII control characters number from 0–31 and are unprintable control codes that are used to control peripherals such as printers, Table 5-3.

The ASCII code of a character is found by combining its Column Number (given in 3-bit binary) with its Row Number (given in 4-bit binary).

The Column Number forms bits 6, 5 and 4 of the ASCII, and the Row Number forms bits 3, 2, 1 and 0 of the ASCII.

Example of use: to get ASCII code for letter “n”, locate it in Column **110**, Row **1110.** Hence its ASCII code is **1101110**.

Table 5-3. ASCII Code

Row Number	**Column Number**							
	000	001	010	011	100	101	110	111
0000	*NUL*	*DLE*	à	0	@	P	`	p
0001	*SOH*	*DC1*	!	1	A	Q	a	q
0010	*STX*	*DC2*	"	2	B	R	b	r
0011	*ETX*	*DC3*	#	3	C	S	c	s
0100	*EOT*	*DC4*	$	4	D	T	d	t
0101	*ENQ*	*NAK*	%	5	E	U	e	u
0110	*ACK*	*SYN*	&	6	F	V	f	v
0111	*BELL*	*ETB*	'	7	G	W	g	w
1000	*BS*	*CAN*	(	8	H	X	h	x
1001	*HT*	*EM*	)	9	I	Y	i	y
1010	*LF*	*SUB*	0	:	J	Z	j	z
1011	*VT*	*ESC*	0	;	K	[	k	{
1100	*FF*	*FS*	,	<	L	\	l	|
1101	*CR*	*GS*	-	=	M	]	m	}
1110	*SO*	*RS*	.	>	N	^	n	~
1111	*SI*	*US*	/	?	O	_	o	DEL

The **Control Code** mnemonics are given in italics above; e.g. *CR* for Carriage Return, *LF* for Line Feed, *BELL* for the Bell, *DEL* for Delete. The Space is ASCII 0100000, and is shown as ◊ here. To write Don Davis in binary, octal, and hexadecimal, refer to Table 5-3.

The text, Don Davis, in binary code is

01000100 01101111 01101110 00100000 01000100 01100001 01110110 01101001 01110011.

The text, Don Davis, in octal code is

104 157 156 040 104 141 166 151 163.

The text, Don Davis, in hexadecimal code is

44 6f 6e 20 44 61 76 69 73.

The binary code is called “machine language” inasmuch as it is the code the computer understands. The compactness that base 8 and base 16 offers over base 2 for use in program editors that then convert these more compact codes into machine language is apparent.

Sound System Engineering in binary code is

01010011 01101111 01110101 01101110 01100100 00100000 01010011 01111001 01110011 01110100 01100101 01101101 00100000 01000101 01101110 01100111 01101001 01101110 01100101 01100101 01110010 01101001 01101110 01100111.

In octal code is

123 157 165 156 144 040 123 171 163 164 145 155 040 105 156 147 151 156 145 145 162 151 156 147.

In hexadecimal code is

53 6f 75 6e 64 20 53 79 73 74 65 6d 20 45 6e 67 69 6e 65 65 72 69 6e 67.

5.7 Bayesian Theory

Bayesian probability theory is sometimes called "common sense, amplified." The rules of probability insure that if two people make the same assumptions and receive the same data, then they will draw identical conclusions. This more general use of probability to quantify beliefs is known as the Bayesian viewpoint. It is also known as the subjective interpretation of probability since the probabilities depend on assumptions. Advocates of a Bayesian approach to data modeling and pattern recognition do not view this subjectivity as a defect, since in their view you cannot do inference without making assumptions.

Laplace, a contemporary of the Reverand Thomas Bayes, gave the first mathematical interpretation of this theory, followed by its being ignored for the next century. In Laplace's interpretation, probability theory is just common sense reduced to numbers, and a probability represents a reasonable degree of belief; not a frequency of occurrence. In the mid-1930s Sir Harold Jeffreys rediscovered the works of Laplace and derived probability theory as an axiomatic theory of inference. E. T. Jaynes, using the methodology of Shannon, the mathematics of Abel, Cox, and the qualitative principles of Laplace, proved that if one represents a reasonable degree of belief as a real number, then the only consistent rules for manipulating probabilities are those given by Laplace. In this wider interpretation of probability theory, called Bayesian probability theory, problems of the form, "what is the best estimate of a parameter one can make from the data and one's prior information" make perfect sense.

Thomas Bayes lived in exciting times, the contemporary of Newton, Berkeley, Laplace, Maclaurin, etc., was a Fellow of the Royal Society. Probability theory as extended logic reproduces many aspects of human mental activity, sometimes in surprising and even disturbing detail. Equations have been found examining the phenomenon of a person who tells the truth and is not believed, even though the disbelievers are reasoning consistently. The theory explains why and under what circumstances this will happen. (I have been the victim of this when the listener heard the facts, but lacked the experience to evaluate facts.)

Jaynes goes on to say:

> *The equations also reproduce a more complicated phenomenon, divergence of opinions. One might expect that open discussion of public issues would tend to bring about a general consensus. On the contrary, we observe repeatedly that when some controversial issue has been discussed vigorously for a few years, society becomes polarized into two opposite extreme camps; it is almost impossible to find anyone who retains a moderate view.*

Probability theory as logic shows how two persons given the same information, may have their opinion driven in opposite directions by it, and what must be done to avoid this.

In such respects, it is clear that probability theory is telling us something about the way our own minds operate when we form intuitive judgments, of which we may not have been consciously aware. Some may feel uncomfortable at these revelations; others may see in them useful tools for psychological, sociological, legal research or artificial intelligence.

It is believed that Sir Harold Jeffreys influenced Alan Turing's work in breaking the enigma code. Jeffreys's book *The Theory of Probability,* published in 1939, would have been accessible to Turing. Since both men were connected to Cambridge University the likelihood that they would have conferred is high.

5.8 Planck System

John Archibald Wheeler (1912 to 2008) was the physicist who developed the system described here, inspired by a similar system by Max Planck.

Planck units represent the smallest possible mass, length, and time possible in our present universe. Some theorists are more willing to speculate than others. But even the boldest acknowledge the "Planck scales" as an ultimate barrier. We cannot measure distances smaller than the Planck length; we cannot distinguish two events, or decide which came first, when the time interval between them is less than the Planck time. These scales are smaller than atoms by just as much as atoms are smaller than stars. There is no prospect of any direct measurement in this domain. It would require particles with energies million, billion times higher than can be produced in the laboratory.

From a practical standpoint, much of the convenience of using Planck units, comes from the fact that the values of the main natural constants, G, H-bar, c, are one when expressed in terms of natural units. To some extent this convenience carries over to several human scale Planck systems in which the natural units have been scaled by powers of ten to make them handier to use.

1. Planck mass 2.1767×10^{-8} kilogram.

2. Planck length 1.6160×10^{-35} meter.
3. Planck time 5.3906×10^{-44} second.
4. Planck frequency is 10^{45} per minute, because it is defined as passing through one radian of phase in each Planck interval of time. As a result the Planck frequency is 10^{40} higher than middle D on the piano.

Planck constant is considered as the fundamental constant of nature.

$$\begin{aligned} h &= 6.6262 \times 10^{-27}\,\text{erg s} \\ &= 6.6262 \times 10^{-34}\,\text{J s} \end{aligned} \tag{5-23}$$

The reduced Planck constant called h-bar is:

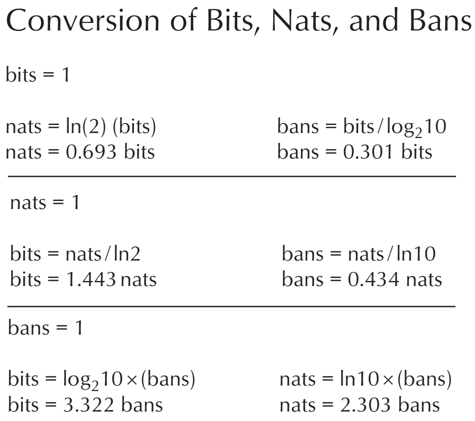
$$h/2\pi \tag{5-24}$$

If, as John Wheeler felt, the universe is "it's from bits," then these parameters are of vital importance.

5.9 Bits, Nats, and Bans

Bits are to the logarithmic base (2), Nats are to the logarithmic base (e), and Bans are to the logarithmic base (10). Bits are used in information work, Nats in physics and Bans have been used in code breaking. In physics, one Nat has an area of two Planck lengths on a side.

The relationship between Bits, Nats, and Bans is shown in Fig. 5-6.



Conversion of Bits, Nats, and Bans

bits = 1

nats = ln(2) (bits)
nats = 0.693 bits

bans = bits/$\log_2 10$
bans = 0.301 bits

nats = 1

bits = nats/ln2
bits = 1.443 nats

bans = nats/ln10
bans = 0.434 nats

bans = 1

bits = $\log_2 10 \times$ (bans)
bits = 3.322 bans

nats = ln10 × (bans)
nats = 2.303 bans

Figure 5-7. Bits, Nats, and Bans.

5.10 A Communication System

Any communication system (acoustic, audio, visual, data, etc.) will consist of a source, message, transmitter, signal path (with the possibility of interfering noise joining the desired signal), a received signal, a receiver, a message, and a destination, Fig. 5-8.

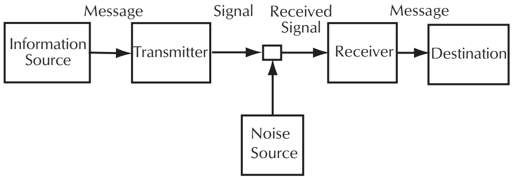


Figure 5-8. A general communications system. (*Credited to Claude Shannon.*)

In a sound reinforcement system, the source will be human (talker, artisan, orchestra, etc.) the message will pass through a transducer into electronics and become an electrical signal. Undesired noise may enter the system from the original acoustic environment or from the electronic path followed. The received signal will again pass through a transducer, i.e., a loudspeaker system or headphones. The message received will then be understood, distorted, or misunderstood.

Our ability to test the system's performance with precision starts at the input transducer and ends at the output transducer. The performer's input and the listener's perception are parameters not readily quantified. Speech intelligibility estimates and speech intelligibility tests of a sufficiently large group of listeners, in the case of talkers, are costly and difficult.

An example of the above was the communication system employed at an airbase in a foreign country to alert personnel to danger. The most expected form of attack was thought to be from the air. The output transducer was a siren. When an unexpected attack occurred, on the ground, by guerrilla warriors, the siren was sounded, but resulted in the wrong actions by the personnel involved on the base.

The system was later replaced by a very high-powered voice system which could issue explicit instructions. Many fire alarm systems share the same problem, such as a large hotel where the fire began in the sound system's central amplifier room, chosen for economic reasons, over a floor-by–floor distributed system. Many survivors could have been saved by a helicopter that landed on the roof, but there was no way to tell the victims to go up to the roof, rather than down the fire stairs.

We hear with regularity the failure of rigging systems installed by unqualified personnel. Overhead canopies collapsing, suspended loudspeaker arrays falling, collapse of staging etc., are often the result of not including all the required professional input at the design stage of a public venue.

All of the above are examples of failure at the design stage to articulate the needs in a language that the architect, engineers, and owners can understand and comprehend. That is the initial communication problem.

5.11 Holography*

Holography was invented in 1947 with the advancement of laser technology. It is a photographic process which does not capture an image of the object being photographed, as in the case with the conventional technique, but rather records the phases and amplitudes of light waves reflected from the object. The wave amplitudes are readily encoded on an ordinary photographic film. The phases are recorded as interference patterns produced by the reflected light and a reference coherent light (from the same laser.) Each point on the hologram received light reflected from every part of the illuminated object, and therefore, contains the complete visual record of the object as a whole. When the hologram obtained from the development of a film exposed in this way is placed in a beam of coherent light, two sets of strong diffracted waves are produced—each an exact replica of the original signal bearing waves that impinged on the plate when the hologram was made. One set of diffracted waves produces a virtual image, which can be seen by looking through the hologram. It appears in a complete three-dimensional form with highly realistic perspective affects. In fact, the reconstructed picture has all the visual properties of the original object.

The holographic and universal information bounds are far beyond the data storage capacities of any current technology and they greatly exceed the density of information on chromosomes and the thermodynamic entropy of water. The maximum allowable information on the surface area can be interpreted as the maximum allowable number of smallish units on each surface. This smallest unit has the size of Planck area as envisioned in quantum theory. According to this theory, the Planck area is a single quantum of space time—it cannot get any smaller than that.

---

* From Universe Review

Bibliography

F. A. Bais and J. D. Farmer, *The Physics of Information*, Cornell University Library.

C. E. Shannon, “The Mathematical Theory of Communication,” *Bell System Technical Journal*.

J. R. Stuart, *Coding High Quality Digital Audio.*

J. R. Stuart, *The Psychoacoustics of Multichannel Audio,* Meridian Audio Lts, Stonehill, Huntingdon, E18 6ED, England.

Chapter 6

Mathematics for Audio Systems

by Don Davis and Eugene Patronis, Jr.

6.1 Engineering Calculations 53
6.2 Precision, Accuracy, and Resolution 54
6.3 Simple Numbers 54
6.4 How to Add Gains and Losses Algebraically 54
6.5 The Factor-Label System 55
  Developing a Conversion Factor 55
  The Foot Pound System 56
  Non-SI Audio Terms 56
  Generating US Dimensions from SI Dimensions 57
  Simplified SI to US and US to SI Conversions 57
6.6 Basic Physical Terms 58
  Conversion Factors from Base Units 59
6.7 Mathematical Operations 60
  Addition 60
  Subtraction 60
  Multiplication 60
  Division 60
  Powers 60
  Roots 60
  Logarithms 60
  Antilogs 61
  Log Multipliers 61
  Antilogs of Multiplied Logarithms 61
  Comparing Arithmetic and Exponential Notation 61
  Comparing Arithmetic, Exponential, and Logarithmic Forms 62
6.8 Complex Number Operations 64
6.9 Decade Calibration 65
6.10 Converting Linear Scales to Logarithmic Scales 66
6.11 Finding the Renard Series for Fractional Octave Spacing 66
  Determining the Number of Octaves in a Given Bandwidth 67
6.12 Radians and Steradians 67
  The Radian 67
  Steradians (Solid Angles) 67
  Symbols 68
  Equations 68
  Complementary Angle for a Given Square Angle and a Desired Arbitrary Angle 68
  Finding Q for a Given Square Angle 69
  Percentage of Spherical Surface Area Covered for a Given Q 69
6.13 Calculating Percentages and Ratios 70
  Decibels and Percentages 71
  Power Changes Downward 71
  Power Changes Upward 71
  Voltage Changes Downward 71
  Voltage Changes Upward 71
  Decibels for Percent Below Reference 72
  Decibels for Percent Above Reference 72
6.14 Useful Math Tables 72
  Scientific Metrology 73

6.15 Angles 75
6.16 A Little Trigonometry 75
6.17 The Origin of the Base of the Natural Logarithm, e 76
6.18 The Complex Plane 77
6.19 Euler's Theorem 77
6.20 Examples 78
    Addition of Complex Numbers. 78
    Subtraction of Complex Numbers. 78
    Products of Complex Numbers 78
    Quotients of Complex Numbers 78
    A Little Digression 78
6.21 Phasors 79
    Addition of Phasors 80
6.22 Rates of Change 80
Bibliography 82

6.1 Engineering Calculations

In my high school days (World War II), I had learned to use simple slide rules. Later at the University level log log duplex decitrig slide rules were a constant companion. At Altec Lansing I used an ancient book of ten place logarithms (Vega) and a Friden mechanical calculator to compute K numbers (the anti-logarithm of a decibel). As I became more involved in acoustic calculations I purchased one of the very first HP 9100s because of its ability to do transcendental functions. This led in the late 1960s to my being selected as one of the guinea pigs for the HP 35 calculator and the opportunity to meet Tom Osborne who had headed the project.

One of my favorite English movies directly after World War II was *The Man in the White Suit*, about an eccentric English inventor who worked his way into large industrial laboratories in the attempt to develop his invention. He would work without charge just to be near the instrumentation that he needed to do his work. In the early 1960s Tom Osborne worked for SCM Corporation (Smith Corona Marchant.)

In the fall of 1963, Osborne told them that he could no longer help them produce a calculator that, in his opinion, was doomed to failure (it was, and it did). He offered to design a machine for them at no cost if they would give him lab space. Later, if they liked what they saw, they could pay him for the time he had spent in the design and construction.

They did not take him up on his offer, but did get excited about his having not given them the plans for the calculator that he must have designed. It seems that they couldn't understand that Osborne could confidently state that he could design a calculator without having already designed it. Much like Mozart, when a client asked him where the score was, Mozart replied, "It's in my head." Osborne and SCM parted company at that point.

Later in that decade after having encountered companies willing to steal his ideas or tie them up so that they never became competitors to them, he encountered Barney Oliver of Hewlett Packard and a remarkable engineering staff capable of not only implementing his ideas but capable of going with him into new realms of technology. The HP 35 (so called because it had thirty-five keys) destroyed slide rules and changed engineering calculations forever.

Bill Hewlett, when asked by the press if he thought the HP 35 would be a success, said he didn't know. When further asked how he justified the expenditure for its creation, he replied that he wanted one. Before it was introduced on the market, analysis by a major consulting firm, had determined that it would fail because of the tiny keys and the reverse Polish notation, (RPN). In Tom Osborne's opinion it succeeded for those very reasons. The HP 35 was so well received that overnight it made the slide rule a relic. Every calculator designed into a computer today came from those pioneer algorithms.

When I showed my HP 35 to Paul Klipsch he asked for the 12th root of 2 (the semitone interval) and with three key punches I showed him 1.0594630944. No more calculations were asked for. Log, trig, square root, etc. tables were retired.

A side note—Hewlett Packard was willing to consider having Altec distribute the HP 35 to the electronic parts industry—Altec owned Allied Radio at that point. I discussed this possibility with the president of Altec, who replied, "We are not in the calculator business," to which I replied, "Neither was HP." It wasn't long before I resigned at Altec and started Synergetic Audio Concepts, Syn-Aud-Con, to teach sound system design to sound contractors.

To quote Richard Heyser:

> *Most of us think of mathematics as those chicken tracks—little wiggle signs. That isn't math! That's the fossil remnants of a thought.*
>
> *The thought is the math. The structured reasoning is the math. And when you start taking things we refer to as common sense and observation to structure that in a reasoning mode—that's math.*
>
> *The axioms and postulates of that which most of us would call common sense is math. When it's dried up and withered and appears as little chicken tracks on a piece of paper, that isn't math! That's just the residue of it, just the shorthand that lets people know that a mind went past here on this page.*
>
> *Math is structured learning.*

Mathematics and physics are the foundation stones that underlie real competence in audio engineering. The mathematical tool most used by audio engineers in describing the physics of sound is the decibel.

A knowledge of mathematics and physics is essential to a realistic understanding of audio engineering. The more background you have in these two disciplines, the greater your ability will be to look past the evidence of the senses to the facts of science. James Moir once told us, "Anything obvious in acoustics is usually wrong." Therefore, we encourage every reader of this book to seek out as much experience in mathematics and physics as is available. This chapter, however, is designed not to review the vast and fascinating horizons available to the scholar but, rather, to outline the minimum requirements that must be reached if you are to know what's going on in an audio system.

The basic tools you must have to understand audio engineering are included in this chapter. With these tools you can make an intelligent choice about the specialized forms of mathematics you might want to study next. We sincerely hope you will study this chapter thoroughly before reading any further inasmuch as the concepts derived here, the terminology developed, and the statements of fundamentals are not repeated.

6.2 Precision, Accuracy, and Resolution

The term "precision" as used in audio engineering is defined as: "The quality of being exactly or sharply defined or stated…which is sometimes indicated by the number of significant digits it contains."

The term "accuracy" as used in audio engineering usually refers to "accuracy rating," which is defined as: The accuracy classification of the instrument. It is given as the limit, usually expressed as a percentage of full-scale value, that errors will not exceed when the instrument is used under reference conditions.

"Resolution" as used in audio engineering is "the act of deriving from a sound, scene, or other form of intelligence a series of discrete elements where from the original may subsequently be synthesized."

You may, for example, have a calculator with a precision to ten digits on which you calculate frequency to a resolution of one-third of an octave using an analyzer with an accuracy of 10%. In such a circumstance, it makes no sense to read out a frequency to ten places on the calculator. One-third octave at 1000 Hz is 230 Hz; plus 10% would be 253 Hz and minus 10% would be 207 Hz. Therefore, you can say with some assurance when tuned to the center of the band, "I'm within 50 Hz of the right answer."

6.3 Simple Numbers

When we use the digits 1, 2, 3, etc., we unconsciously assume that:

1. Numbers have magnitude—Two is larger than one; three is larger than two, etc.
2. Numbers have signs—If we wrote down our assumption, we would write +1, +2, +3, etc.
3. Numbers are ratios—We recognize this every time we do a decibel problem: 1/1, 2/1, 3/1, etc.
4. Numbers have exponents—We don't always indicate them, but we assume them. For example, we could write 1^1, 2^1, 3^1, etc.
5. Zero has no magnitude, ratio, exponent, or sign. Zero is the symbol for the idea of nothingness. It is the placeholder in large numbers.
6. Any digit raised to the exponent of zero equals unity—For example, 10^0, 100^0, 1000^0 equal 1.
7. Digits raised to positive number exponents produce larger numbers: $2^2 = 4$.
8. Digits raised to fractional number exponents are roots of numbers and produce smaller numbers: $2^{1/2} = \sqrt{2} = 1.414\ldots$.
9. Digits raised to negative number exponents are fractional numbers:

   $$2^{-2} = 1/4\ ,\ (2/1)^{-2} = (0.25/1)^1 = 0.25\ .$$

10. Numbers are reciprocals:

    $$1 = (1/1)\ ,\ 2 = (1/0.5)\ ,\ 3 = 1/0.33\ .$$

11. Numbers are roots:

    $$1 = \sqrt{1}\ ,\ 2 = \sqrt{4}\ ,\ 3 = \sqrt{9}\ .$$

When labels are added to numbers, they can no longer be thought of as simple. Yet even in their simple forms the meanings they have can easily be overlooked by the casual observer. The preceding list certainly does not exhaust the assumptions you can make, but it does suggest how simple numbers are perceived.

6.4 How to Add Gains and Losses Algebraically

The mathematical signs plus (+) and minus (–) are directional indicators. Imagine a line with zero at its center. Negative numbers proceed from zero to the left and positive numbers proceed to the right of the same zero point.

To sum the following gains and losses:

–148 + 100 + 47.5 + 48 + 56.5 – 10.5, the following procedure is suggested:

Add all the negative numbers

$$\begin{array}{r} -148.0 \\ -10.5 \\ \hline -158.5 \end{array}$$

Add all the positive numbers

$$\begin{array}{r} +100.0 \\ +47.5 \\ +48.0 \\ +56.5 \\ \hline +252.0 \end{array}$$

Sum the two results:

$$\begin{array}{r} +252.0 \\ -158.5 \\ \hline +93.5 \end{array}$$

This procedure is easy. Most systems exhibit gain (positive numbers are the largest value). Where a system exhibits loss (negative numbers are the largest value).

To sum the following gains and losses: $-148 + 100 + 40 - 60$ use the following procedure:

1. Sum all of the negative numbers:

$$\begin{array}{r} -148 \\ -60 \\ \hline -208 \end{array}$$

2. Sum all of the positive numbers:

$$\begin{array}{r} +100 \\ +40 \\ \hline +140 \end{array}$$

3. Determine whether the negative numbers are larger in value than the positive numbers: $208 > 140$.
4. Change the signs of the negative to positive and the positive to negative, and subtract the smaller value from the larger value:

$$\begin{array}{r} +208 \\ -140 \\ \hline +68 \end{array}$$

5. Reverse the sign of the answer—The reverse sign of +68 is –68.

6.5 The Factor-Label System

The factor-label system is most useful. Factors are numbers that represent quantities of something or a measure of something on a measurement scale. Labels are the actual units or titles used on our measurement scale. When combined, a factor-label results. Given 2 apples, “2” is the factor and “apples” is the label, or 60°F, “60” is the position on the measurement scale, and “F” is the label.

Careless use of labels has led to such statements as “It’s a watt,” when what is meant is “It’s one watt.” In this example, the factor is implied when the label is used.

6.5.1 Developing a Conversion Factor

Suppose we had a statement that we are at a velocity of 88 feet per second (88 ft/s). This can be written as:

$$V = \frac{88\text{ ft}}{1\text{ s}}$$

where,
“88” and “1.0” are factors,
“ft” and “s” are labels.

Velocity in miles per hour (mi/h), can be written as:

$$V = \frac{x\text{ mi}}{1.0\text{ h}} \tag{6-1}$$

Change the labels by developing a conversion factor as follows:

$$\frac{x\text{ mi}}{1.0\text{ h}} = \frac{88\text{ ft}}{1.0\text{ s}} \times \frac{3600\text{ s}}{1.0\text{ h}} \times \frac{1.0\text{ mi}}{5280\text{ ft}}$$

where,

$$\frac{3600\text{ s}}{1.0\text{ h}} = 1\,,$$

$$\frac{1.0\text{ mi}}{5280\text{ ft}} = 1\,.$$

Cancel labels to get the correct factor relationship for the labels that remain:

$$\frac{88\,\cancel{\text{ft}}}{1\,\cancel{\text{s}}} \times \frac{3600\,\cancel{\text{s}}}{1\text{ h}} \times \frac{1.0\text{ mi}}{5280\,\cancel{\text{ft}}}$$

so that

$$\frac{(88)(3600)}{5280\text{ h}} = \frac{60\text{ mi}}{1.0\text{ h}}$$

To find the conversion factor, take the factors employed to obtain the new labels as the conversion factor (3600/5280 = 0.6818818…). The original factor, (88), is not used to generate the conversion factor because it is what we want to convert:

88 (0.681818…) = 60

We can now write in a table of conversion factors.

To Find	Multiply	By
MPH (mi/h)	ft/s	0.681818

The reciprocal of 0.681818 or (1/0.681818) will do the inverse (change mi/h to ft/s).

The orderliness of this technique, especially in programming computers and calculators, recommends its use by audio engineers.

In this book we have chosen to use those labels most conveniently at hand without undue regard to various labeling systems. The two principal labeling systems in current use are:

1. The United States system (US), formerly called the English system.
2. The Systeme International d'Unites (SI).

To our best knowledge, no country in the world is exclusively on either system. Consequently, you must know how to operate any system encountered in the world of practical engineering.

6.5.2 The Foot Pound System

In parallel with the development of the CGS system came what was seen as its Imperial equivalent, the foot-pound-second system proposed by W. Stroud in 1880. It became very widely employed in all branches of engineering, and most technical papers written in Britain, the USA, and other parts of the English-speaking world before about 1960 would have used these units, although scientific papers tended to use CGS units.

Its popularity in engineering was due not only to the use of Imperial units as its base, but also because the pound and the foot were felt to be more convenient for engineers than the too-small centimeter and gram, and the too-large meter and kilogram. Although it was strictly speaking, a non-decimal system, this was technically irrelevant since quantities could be expressed in decimals of feet, pounds, etc., so that the criticism of complexity and calculations usually aimed at the Imperial system did not necessarily apply.

Its main problem was that the pound had long been in common use as a unit of both weight and mass. This makes no difference in general and commercial usage, since, because of the Earth's gravity, a mass of 1 pound weighs exactly 1 pound. Because the moon's gravitational force is about one-sixth that of the Earth, the 1 pound mass would weigh only one-sixth of a pound, although the mass itself would not have changed. Weight, therefore, is the force with which a mass is attracted by gravity, and, since it is an entirely different quantity, it requires a different unit.

However, because in general use the pound had always been appreciated as a unit of weight, there was a tendency among engineers to continue to use it in this way. Therefore in the United States a variant of the FPS system usually termed, technical, gravitational, or engineer's units, the pound-force (lbf) was taken as a base unit, and the unit of mass was derived from it. This unit was named the slug, and was the mass which when acted upon by one pound force experienced an acceleration of $1\ \text{ft}/\text{s}^2$, so it was equal to 32.17 pounds.

It is worth noting that there is nothing inherently inconsistent in a system based on the foot and the pound. Decimalized, with a single set of force and mass units, and integrated with electrical and molar quantities, it could've been just as consistent and international as the metric-based SI.

There is a feeling among human beings that units spaced on the human body are somehow more comprehensible than those derived from the circumference of the earth, or referred to the energy level of an atom. Today's engineers need to be thoroughly conversant with any system that provides usefully dimensioned units. A case in point, the use of Planck units in information theory.

6.5.3 Non-SI Audio Terms

There is a vast list of units outside the accepted SI (System International) list that may be used because of their obvious utility. These are listed in a series of addendum in the SI standards. Two key ones regard digital terms and level terms:

1. The dB is used rather than arithmetic ratios or percentages because when circuits are connected in tandem, expressions of power level, in dB, may be arithmetically added and subtracted. For example, in an optical link, if a known amount of optical power in dBm is launched into a fiber, and the losses, in dB, of each component (e.g., connectors, splices, and lengths of fiber) are known, the overall link loss may be quickly calculated with simple addition and subtraction.
2. The neper is often used to express voltage and current ratios, whereas the decibel is usually used to express power ratios.

The above notes are frequently ignored, but they do represent notes in the SI standards. The now prevalent use of voltage and current ratios expressed as levels in decibels and its carryover in some standards, in order to placate technicians, is in my opinion, a mistake.

The digital situation is even messier than the level situation as evidenced by the quotes below written by Bruce Barrow for the IEEE.

> *Once upon a time, computer professionals noticed that 2^{10} was very nearly equal to 1000 and started using the SI prefix 'kilo' to mean 1024. That worked well enough for a decade or two because everybody who talked kilobytes knew that the term implied 1024 bytes. But, almost overnight a much more numerous 'everybody' bought computers, and the trade computer professionals needed to talk to physicists and engineers and even to ordinary people, most of whom know that a kilometer is 1000 m and a kilogram is 1000 g.*
>
> *Then data storage for gigabytes, and even terabytes, became practical, and the storage devices were not constructed on binary trees, which meant that for many practical purposes, binary arithmetic was less convenient than decimal arithmetic. The result is that today 'everybody' does not 'know' what a megabyte is. When discussing computer memory, most manufacturers use megabyte to mean 2^{20} =1 048 576 bytes, but the manufacturers of computer storage devices usually use the term to mean 1 000 000 bytes. Some designers of local area networks have used megabit per second to mean 1 048 576 bits per second, but all telecommunications engineers use it to mean 10^6 bits per second. And as if two definitions of the megabyte are not enough, a third megabyte of 1 024 000 bytes is the megabyte used to format the familiar 90 mm (3 ½ inch), '1.44 MB' diskette. The confusion is real, as is the potential for incompatibility in standards and in implemented system.*
>
> *Faced with this reality, the IEEE standards Board decided that IEEE standards will use the conventional, internationally adopted definitions of the SI prefixes. Mega will mean 1 000 000, except that the base 2- definitions may be used (if such usage is explicitly pointed out on a case - by - case basis) until such time that prefixes for binary multiples are adopted by an appropriate standards body.*

The NIST reference on constants, units, and uncertainty with regard to the international system of units, should be required reading for any engineer.

6.5.4 Generating US Dimensions from SI Dimensions

If we remember that there are 2.54 cm/1.0 in, it is simple to convert SI linear dimension to the US system.

Example

Suppose we are in Europe and traveling with a friend in his Porsche; we notice that he cruises the Autobahn at 230 kilometers per hour (230 km/h). (SI dimensions are not the only difference between Americans and Europeans.) This seems a little fast, so you check it. The conversion factor from kilometers per hour (km/h) to miles per hour (mi/h) is:

$$\frac{100{,}000\text{ cm}}{1\text{ km}} \times \frac{1\text{ mi}}{5280\text{ ft}} \times \frac{1\text{ ft}}{12\text{ in}} \times \frac{1\text{ in}}{2.54\text{ cm}} = \frac{0.62\text{ mi}}{1\text{ km}}$$

which means that to convert km/h to mi/h, multiply km/h by 0.62137…. To convert mi/h to km/h, take the reciprocal (1⁄*x*) of the SI to US conversion factor (0.6213) to obtain a US to SI conversion factor: $1/0.6213 = 1.60934$, which, for most purposes, can be rounded to 1.61 km/h to each 1.0 mi ⁄h.

To feel comfortable with 230 km/h (i.e., 143 mph), you don't need to be an expert in SI, but an expert driver. The same can be said about audio and acoustic measurements. It matters little which system you use or how expert you are with it if you mismeasure the audio or acoustic parameter under test.

6.5.5 Simplified SI to US and US to SI Conversions

When making conversions, use the following logic. The statement defines the answer desired:

Statement Factor Label

$$\frac{\text{mi}}{\text{km}} = \frac{1.0\,\text{mi}}{5280\,\cancel{\text{ft}}} \times \frac{1.0\,\cancel{\text{ft}}}{12\,\cancel{\text{in}}} \times \frac{1.0\,\cancel{\text{in}}}{2.54\,\cancel{\text{cm}}} \times \frac{100{,}000\,\cancel{\text{cm}}}{1.0\,\text{km}}$$

$$= \frac{0.621\,\text{mi}}{\text{km}}$$

All factor-labels have to equal unity. They can be any set of dimensions so long as they equal 1.0.

They should be arranged so that the cancelling labels are adjacent and the only labels left after the cancellations are the two that were in the statement. At that point all the factors are multiplied and divided to give the desired answer. In the case above, the statement is "miles per kilometer." The answer is 0.621 mi/km.

It could have been mi/km for 100 km.

Statement Factor Label

$$\frac{\text{mi}}{\text{km}} = \frac{1.0\,\text{mi}}{5280\,\cancel{\text{ft}}} \times \frac{1.0\,\cancel{\text{ft}}}{12\,\cancel{\text{in}}} \times \frac{1.0\,\cancel{\text{in}}}{2.54\,\cancel{\text{cm}}} \times \frac{10{,}000{,}000\,\cancel{\text{cm}}}{100\,\text{km}}$$

$$= \frac{0.621\,\text{mi}}{\text{km}}$$

1. Meters to feet (m to ft)

$$\frac{\text{ft}}{\text{m}} = \frac{1\text{ in}}{2.54\text{ cm}} \times \frac{100\text{ cm}}{1\text{ m}} \times \frac{1\text{ ft}}{12\text{ in}} \qquad (6\text{-}2)$$

2. Feet to meters (ft to m)

$$\frac{\text{m}}{\text{ft}} = \frac{2.54\text{ cm}}{1\text{ in}} \times \frac{1\text{ m}}{100\text{ cm}} \times \frac{12\text{ in}}{1\text{ ft}} \qquad (6\text{-}3)$$

3. Square meters to square feet (m^2 to ft^2)

$$\frac{\text{ft}^2}{\text{m}^2} = \left[\frac{1\text{ in}}{2.54\text{ cm}} \times \frac{100\text{ cm}}{1\text{ m}} \times \frac{1\text{ ft}}{12\text{ in}}\right]^2 \qquad (6\text{-}4)$$

4. Square feet to square meters (ft^2 to m^2)

$$\frac{\text{m}^2}{\text{ft}^2} = \left[\frac{2.54\text{ cm}}{1\text{ in}} \times \frac{1\text{ m}}{100\text{ cm}} \times \frac{12\text{ in}}{1\text{ ft}}\right]^2 \qquad (6\text{-}5)$$

5. Cubic meters to cubic feet (m^3 to ft^3)

$$\frac{\text{ft}^3}{\text{m}^3} = \left[\frac{1\text{ in}}{2.54\text{ cm}} \times \frac{100\text{ cm}}{1\text{ m}} \times \frac{1\text{ ft}}{12\text{ in}}\right]^3 \qquad (6\text{-}6)$$

6. Cubic feet to cubic meters (ft^3 to m^3)

$$\frac{\text{m}^3}{\text{ft}^3} = \left[\frac{2.54\text{ cm}}{1\text{ in}} \times \frac{1\text{ m}}{100\text{ cm}} \times \frac{12\text{ in}}{1\text{ ft}}\right]^3 \qquad (6\text{-}7)$$

The SI base units and their US counterparts are seen in Table 1-1 along with a listing of derived units of interest to sound system engineers in Table 1-2.

Decimal notation can be used with either system and is not an exclusive feature of either. Time has not been decimalized. Universal physical constants continue to defy humanly contrived measurements in search of a coherent system.

The Planck system shows promise based on the Planck second where:

1. Planck second = 10^{-44} s.
2. Planck length = 10^{-33} cm.
3. Planck mass = 10^{-6} gm.

6.6 Basic Physical Terms

Force. *Force* is described as one kilogram moved one meter per second per second ($1\,\text{kg(m)/s}^2$) which is equal to one newton:

$$F = MA \qquad (6\text{-}8)$$

where,
F is the force in newtons,
M is mass in kilograms,
A is acceleration in meters per second per second.

$$\text{N} = \text{kg} \times \text{m} \times \text{s}^{-2} \qquad (6\text{-}9)$$

Pressure. *Pressure* is the amount of force that is acting on each unit area. A pressure of one newton per square meter (N/m^2) equals one pascal (Pa):

$$\begin{aligned}\frac{\text{N}}{\text{m}^2} &= \text{Pa}\\ &= \frac{\text{kg}}{\text{m} \times \text{s}^2} \qquad (6\text{-}10)\\ &= \text{m}^{-1} \times \text{kg} \times \text{s}^{-2}\end{aligned}$$

Work and Energy. When we do work we expend energy. The energy that a mass has, as a result of its position, is called *potential energy*. The energy that a mass has, as a result of its motion, is called *kinetic energy*. Work (in joules) equals force (in newtons) times distance moved in the direction of the force (in meters):

$$\begin{aligned}W &= F \times D\\ &= \frac{\text{kg} \times \text{m}^2}{\text{s}^2} \qquad (6\text{-}11)\\ &= \text{m}^2 \times \text{kg} \times \text{s}^{-2}\end{aligned}$$

Power. *Power* is the rate at which work is done. The average power (in watts) is equal to the work (in joules) divided by the time in seconds:

Table 6-1. Commonly Used Conversions

Quantity	Base SI unit	Symbol	Relationship	Base US unit	Symbol
Length	Meter	m	m = 3.281 ft	Feet	ft
Mass	Kilogram	kg	kg = 0.06852 slg	Slug	slg
Time	Second	s	s = s	Second	s
Electric current	Ampere	A	A = A	Ampere	A
Thermodynamic temperature	Kelvin	K	K = (°F + 459.67)/1.8	Degree Fahrenheit	°F
Amount of Substance	Mole	mol	mol = mol	Mole	mol
Luminous intensity	Candela	cd	cd =cd	Candela	cd
Plane angle	Radian	rad	2π rad = 360°	Degree	∠°
Solid angle	Steradian	sr	4πsr = sphere	Sphere	sph

$$\begin{aligned} P &= \frac{J}{s} \\ &= \frac{\mathrm{kg} \times \mathrm{m}^2}{\mathrm{s}^3} \\ &= \mathrm{m}^2 \times \mathrm{kg} \times \mathrm{s}^{-3} \end{aligned} \tag{6-12}$$

$$(746\ \mathrm{W} = 1\ \mathrm{hp} = 58.73\ \mathrm{dBm})$$

Table 6-2. Common Electrical and Acoustical Derived Units

Quantity	Unit	Symbol
Pressure	Pascal	$\mathrm{Pa} = \mathrm{kg} \cdot \mathrm{m}^{-1} \cdot \mathrm{s}^{-2}$
Power	Watt	$\mathrm{W} = \mathrm{m}^2 \cdot \mathrm{kg} \cdot \mathrm{s}^{-3}$
Force	Newton	$\mathrm{N} = \mathrm{m} \cdot \mathrm{kg} \cdot \mathrm{s}^{-2}$
Work	Joule	$\mathrm{J} = \mathrm{m}^2 \cdot \mathrm{kg} \cdot \mathrm{s}^{-2}$
Electromotive force	Volt	$\mathrm{V} = \mathrm{m}^2 \cdot \mathrm{kg} \cdot \mathrm{s}^{-3} \cdot \mathrm{A}^{-1}$
Electrical resistance	Ohm	$\Omega = \mathrm{m}^2 \cdot \mathrm{kg} \cdot \mathrm{s}^{-3} \cdot \mathrm{A}^{-2}$
Frequency	Hertz	$\mathrm{Hz} = \mathrm{s}^{-1}$
Capacitance	Farad	$\mathrm{F} = \mathrm{m}^{-2} \cdot \mathrm{kg}^{-1} \cdot \mathrm{s}^4 \cdot \mathrm{A}^2$
Inductance	Henry	$\mathrm{H} = \mathrm{m}^2 \cdot \mathrm{kg} \cdot \mathrm{s}^{-2} \cdot \mathrm{A}^{-2}$

6.6.1 Conversion Factors from Base Units

To generate a factor for converting pounds force per square foot (lbf / ft^2 into pascals (Pa):

$$\begin{aligned} N &= \frac{\mathrm{m} \times \mathrm{kg}}{\mathrm{s}^2} \\ &= (\mathrm{m} \times \mathrm{kg} \times \mathrm{s}^{-2}) \end{aligned} \tag{6-13}$$

The pascal is a newton per square meter (N/m^2), therefore,

$$Pa = \mathrm{m} \times \mathrm{kg} \times \mathrm{s}^{-2} \times \mathrm{m}^{-2} \tag{6-14}$$

or

$$Pa = \frac{\mathrm{kg}}{\mathrm{m} \times \mathrm{s}^2} \tag{6-15}$$

A pound of force (lbf) will accelerate a slug of mass one foot per second per second. Therefore a pound of force per square foot is given by:

$$\frac{\mathrm{lbf}}{\mathrm{ft}^2} = \frac{\mathrm{slug}}{\mathrm{ft} \times \mathrm{s}^2} \tag{6-16}$$

From Table 6-1, a kilogram is equivalent to 0.06852 slug. This means there are 1/0.06852 kilograms per slug. Similarly, there is 0.3048 meter per foot. Substituting into Eq. 6-16

$$\begin{aligned} \frac{\mathrm{lbf}}{\mathrm{ft}^2} &= \frac{\mathrm{slug}}{\mathrm{ft} \times \mathrm{s}^2} \times \frac{14.594\ \mathrm{kg}}{\mathrm{slug}} \times \frac{\mathrm{ft}}{0.3048\ \mathrm{m}} \\ &= 47.880 \frac{\mathrm{kg}}{\mathrm{m} \times \mathrm{s}^2} \end{aligned}$$

The conclusion is

$$\frac{\mathrm{lbf}}{\mathrm{ft}^2} = 47.880260\ \mathrm{Pa} \tag{6-17}$$

Atmospheric pressure is 2116.2 lb/ft^2 at the earth's surface (sea level). Therefore 47.880260 Pa × 2116.2 lb/ft^2 produces 101,324.2055 Pa. Our zero reference sound pressure is 0.00002 Pa (20 µPa), so full modulation of atmospheric pressure would be a sound pressure level of

$$20 \log \left[\frac{101{,}324.2055\ \mathrm{Pa}}{0.00002\ \mathrm{Pa}} \right] = 194.0937\ \mathrm{dB}$$

6.7 Mathematical Operations

Rule 1

If a number has a (+) sign, starting at zero, go in a straight line along the x-axis toward 0°, Fig. 6-1. If a number has a (–) sign, starting at zero go along the x-axis toward 180°. (Regard a (–) sign as an instruction to revolve the sign assignment 180°.)

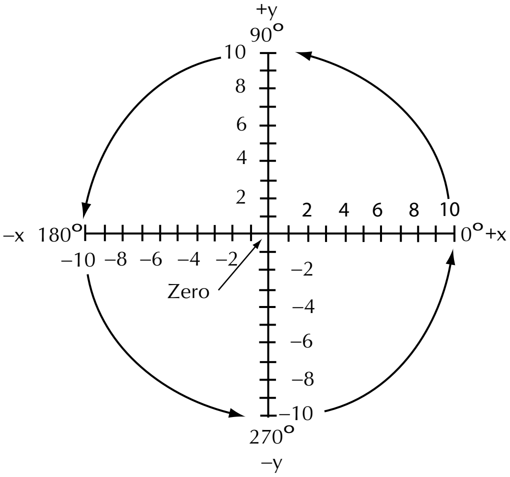


Figure 6-1. Basic notation.

The length of the line toward either 0° or 180° is determined by the magnitude of the number. For example, +4 goes farther along the 0° axis line than +2. All + signs add in magnitude toward 0°. All – signs add in magnitude toward 180°. When minus signs are added to plus signs, the number having the greatest magnitude determines the sign.

Rule 2

Multiplying and dividing magnitudes follow the same rules with the following variations. Every minus sign encountered rotates the magnitude sign 180°. Therefore, +2 × (–2) = (2 × 2) + 180°, or –4. But (–2) × (–2) = 2 × 2 + 360° or +4. Remember every minus sign rotates the sign assignment 180°; therefore, (–2) × (–2) × (–2) = (2 × 2 × 2) + 540°, or –8. This is due to the fact that the symbol "–" is both an operator and a sign.

6.7.1 Addition

Adding is taking the sum of two numbers (1 + 1 = 2) or counting the total number.

6.7.2 Subtraction

Subtracting is changing the direction of addition from 0° to 180°. As shown in Fig. 6-2, +2 –3 means we first move two marks in the positive direction from 0 to +2. We then move three marks in the negative direction from that mark (+2). This results in a final position on the marks of –1.

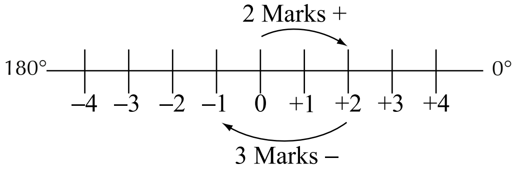


Figure 6-2. Directions for addition and subtraction.

6.7.3 Multiplication

Multiplication is a form of repeated addition. The notations 3 × 6, 3 · 6, and (3)(6) all mean add three sixes together: 6 + 6 + 6 = 18.

6.7.4 Division

Division is a form of repeated subtraction. The notations 6 ÷ 3, 6/3, and $\frac{6}{3}$ all mean find the number that when subtracted from 6 three times results in zero: +6 – 2 – 2 – 2 = 0.

6.7.5 Powers

Taking numbers to higher powers (exponents) is a form of repeated multiplication. Both 2^4, and 2 exp 4 mean multiply 2 by itself four times: 2 × 2 × 2 × 2 = 16.

6.7.6 Roots

Roots are a form of repeated division. The notations $16^{1/4}$, $\sqrt[4]{16}$, and 16 exp 0.25 all mean find the number that can be divided into 16 four times with a remainder of zero. In this example the answer is 2.

6.7.7 Logarithms

Taking the logarithm of a number is a method of expressing that number as an exponent of some chosen base number. Using the base 10 allows simple illustrations to be formed.

$$10 \text{ x } 10 = 100 = 10^2$$

$$\log_{\underset{\text{base}}{10}} \underset{\text{number}}{100} = \underset{\text{logarithm}}{2}$$

Logarithm operation means take the log of the number to the base.

6.7.8 Antilogs

The inverse of this operation is the number ratio expressed as an exponent of the base. Both $\log^{-1}$, and antilog, mean:

$$\text{antilog}_{10} 2 = 10^2 \tag{6-18}$$

6.7.9 Log Multipliers

Logarithmic multipliers are used in audio and acoustics such that

$$M \log_b \frac{a}{c} = NM \tag{6-19}$$

where,

M is the multiplier,

$\log_b (a/c)^M$ is identical to $M \log_b a/c$,

b is the base (it may be any value other than zero or unity),

a/c is the ratio being converted into a logarithm,

NM is the logarithm times the multiplier.

6.7.10 Antilogs of Multiplied Logarithms

To find the antilog of multiplied logarithms, the first step is to remove the multiplier so the quantity can be treated as a normal logarithm. The inverse operation is, of course, division so that $NM/M = N$ is obtained. Then the antilog is found by the standard method $b^N = a/c$. These two valuable tools are written as:

$$M \log_b \frac{a}{c} = NM \qquad \text{(Log form)} \tag{6-20}$$

$$\text{antilog}_b \frac{a}{c} = b^{\frac{NM}{M}} \qquad \text{(Antilog form)} \tag{6-21}$$

6.7.11 Comparing Arithmetic and Exponential Notation

Arithmetic and exponential notation are compared in Table 6-3. Note that positive exponents move the decimal point to the right. Negative exponents move the decimal point to the left of the first numeral by a number equal to the exponent. Also note that roots are indicated by fractional exponents. For example: $10^{0.5} \times 10^{0.5} = 10^{(0.5 + 0.5)} = 10$; therefore, $10^{0.5} = \sqrt{10}$.

Table 6-3. Comparison of Arithmetic and Exponential Notation

Arithmetic Notation	Exponential Notation	Result
10×10	10^2	100
$10 \times 10 \times 10$	10^3	1000
$10 \times 10 \times 10 \times 10$	10^4	10,000
$10 \times 10 \times 10 \times 10 \times 10$	10^5	100,000
$100{,}000/10$	10^4	10,000
$10{,}000/10$	10^3	1000
$(1000)/10$	10^2	100
$100/10$	10^1	10
$10/10$	10^0	1
$1/10$	10^{-1}	0.1
$0.1/10$	10^{-2}	0.01
$(0.01)/10$	10^{-3}	0.001
$0.001/10$	10^{-4}	0.0001
$0.0001/10$	10^{-5}	0.00001
100×1000	$10^2 \times 10^3$ or $10^{(2+3)}$	100,000
10×100	$10^1 \times 10^2$ or $10^{(1+2)}$	1000
$100{,}000/1000$	$10^5/10^3$ or $10^{(5-3)}$	100
$1000/100$	$10^3/10^2$ or $10^{(3-2)}$	10
$\sqrt{10}$	$10^{0.5}$ or $10^{1/2}$	3.162
$\sqrt[3]{10}$	$10^{0.33}$ or $10^{1/3}$	2.154
$\sqrt[4]{10}$	$10^{0.25}$ or $10^{1/4}$	1.778
$\sqrt[5]{10}$	$10^{0.2}$ or $10^{1/5}$	1.585

General Rule

Any positive real ratio a/c can be expressed by two numbers, b and n, in exponential form:

$$\frac{a}{c} = b^n \tag{6-22}$$

Either b or n can be chosen arbitrarily, with certain obvious restrictions, but the choice of one determines the other. If a value of 10 is assigned to b and $a = 10$ and $c = 1$, $10/1 = 10^1$ consequently, $n = 1$.

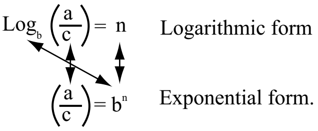
Another way to express the same quantities is in the logarithmic form:

$$\begin{aligned} \log_b\frac{a}{c} &= \log_b b \times n \\ \therefore \frac{\log_b\frac{a}{c}}{\log_b b} &= n \end{aligned} \tag{6-23}$$

or

$$\begin{aligned} \log_b b &= 1 \\ \therefore \log_b\frac{a}{c} &= n \end{aligned} \tag{6-24}$$

Each element in the logarithmic and exponential forms has an equivalent element in the other form:



The arrows indicate how the quantities transpose from one form to the other. The logarithmic form is read, "log to the base b of a over c equals n." The exponential form is read, "a over c equals the base b raised to the power n."

6.7.12 Comparing Arithmetic, Exponential, and Logarithmic Forms

Since the logarithm of a number is the exponent that the base must be raised to in order to equal that number, we can write logarithmic equivalents of arithmetic equations by remembering how exponents are manipulated, see Table 6-4. Table 16-5 shows the terms and symbols for logarithmic scales. Once we recognize that logarithms are the exponents of some base and that they follow exponential rules of manipulation, we will understand decibel notation, which also uses a logarithmic system of notation. Practice with these very basic concepts can serve an audio professional well.

Rule

The reciprocal of an exponent of a base may be used as the root of the base without a change in value (of the antilog):

$$10^{0.5} = \sqrt{10}$$

Table 6-4. Equivalent Arithmetic, Exponential, and Logarithmic Forms

Arithmetic	Exponential	Logarithmic
$A \cdot B$	$A^1 \cdot B^1$	$10^{(\log A + \log B)}$
$2 \cdot 3$	$2^1 \cdot 3^1$	$10^{(\log 2 + \log 3)}$
A/B	A^1/B^1	$10^{(\log A - \log B)}$
$4/2$	$4^1/2^1$	$10^{(\log 4 - \log 2)}$
A^x	A^x	$10^{x \log A}$
2^3	2^3	$10^{3 \log 2}$
$\sqrt[x]{A}$	$A^{1/x}$	$10^{\frac{\log A}{x}}$
$\sqrt[2]{3}$	$3^{1/2}$	$10^{\frac{\log 3}{2}}$
$A^x \cdot A^y$	$A^{(x+y)}$	$10^{(x \log A + y \log A)}$
$3^2 \cdot 3^3$	$3^{(2+3)}$	$10^{(2 \log 3 + 3 \log 3)}$
$(A \cdot B)^x$	$A^x \cdot B^x$	$10^{(x \log A + x \log B)}$
$(2 \cdot 3)^4$	$2^4 \cdot 3^4$	$10^{(4 \log 2 + 4 \log 3)}$
$\sqrt[x]{A/B}$	$(\sqrt[x]{A})/(\sqrt[x]{B})$	$10^{\left(\frac{\log A}{x} - \frac{\log b}{x}\right)}$
$\sqrt[4]{3/2}$	$(\sqrt[4]{3})/(\sqrt[4]{2})$	$10^{\left(\frac{\log 3}{4} - \frac{\log 2}{4}\right)}$
$(A^x)^y$	A^{xy}	$10^{xy \log A}$
$(2^3)^4$	$2^{(3 \times 4)}$	$10^{3 \times 4 \log 2}$
$\sqrt[x]{A \cdot B}$	$\sqrt[x]{A} \cdot \sqrt[x]{B}$	$10^{\left(\frac{\log A}{x} + \frac{\log b}{x}\right)}$
$\sqrt[4]{2 \cdot 3}$	$\sqrt[4]{2} \cdot \sqrt[4]{3}$	$10^{\left(\frac{\log 3}{4} + \frac{\log 2}{4}\right)}$
$A^{1/x}$	$\sqrt[x]{A}$	$10^{\left(\frac{\log A}{x}\right)}$
$2^{1/3}$	$\sqrt[3]{2}$	$10^{\left(\frac{\log 2}{3}\right)}$
A^x/A^y	$A^{(x-y)}$	$10^{(x-y)\log A}$
$2^4/2^3$	$2^{(4-3)}$	$10^{(4-3)\log 2}$
$(A/B)^x$	A^x/B^x	$10^{x \log A - x \log B}$

Table 6-4. (cont.) Equivalent Arithmetic, Exponential, and Logarithmic Forms

Arithmetic	Exponential	Logarithmic
$(3/2)^4$	$3^4/2^4$	$10^{4\log 3 - 4\log 2}$
A^{-x}	$1/A^x$	$10^{-x\log A}$
2^{-4}	$1/2^4$	$10^{-4\log 2}$
$\sqrt[x]{\sqrt[y]{A}}$	$\sqrt[xy]{A}$	$10^{\left(\frac{\log A}{xy}\right)}$
$\sqrt[2]{\sqrt[3]{4}}$	$\sqrt[2\times 3]{4}$	$10^{\left(\frac{\log 4}{2\times 3}\right)}$
$A^{x/y}$	$\sqrt[y]{A^x}$	$10^{\left(\frac{x\log A}{y}\right)}$
$2^{3/4}$	$\sqrt[4]{2^3}$	$10^{\left(\frac{3\log 2}{4}\right)}$
A^{x^y}	$A^{(x^y)}$	$10^{10^{(\log\log A + y\log x)}}$
2^{3^4}	$2^{(3^4)}$	$10^{10^{(\log\log 2 + 4\log 3)}}$
A^0	1	$10^{0\log A}$

Table 6-5. Terms and Symbols for Logarithmic Scales*

Physical Quantity	Base	Name of One Order	Symbol
Power attenuation or gain	10	Bel	B
Stellar magnitude (Brightness^{-1})	$100^{1/5} = 2.512$	Magnitude	
Musical pitch and other harmonic analysis (frequency)	$(f_H/f_L = 2^n)$, where f_L is the lower frequency, f_H is the higher frequency, and n is the number of octaves.	Octave	OC
Photographic exposure settings	$10^{3/10} = 1.995$	Step or Stop	ST
Various electrical, acoustic, and mechanical (proposed for general use)	e = 2.718	Neper	Np or ln
Proposed for general use	10	Brigg	Br
Proposed for general use	b	Order to base b	ORD$_b$

* Proposed by Calvin S. McCamy, N.B.S.

Rule

The reciprocal of a root of a base may be used as the exponent of the base without a change in value (of the antilog):

$$\sqrt[2]{10} = 10^{0.5}$$

Rule

Exponents of a base, multiplied together, are equivalent to the antilog of one exponent raised to the second exponent:

$$10^{2\times 3} = 100^3 = 1000^2$$

Rule

One exponent of a base, divided by a second exponent, is equivalent to the antilog of the first exponent taken to the root of the second exponent:

$$10^{4/2} = \sqrt[2]{10{,}000}$$

Rule

One exponent of a base, divided by a second exponent, is the equivalent to the second exponent taken as the root of the base, raised to the first exponent:

$$10^{4/2} = \sqrt[2]{10^4}$$

Example

Find the decimal equivalent of $10^{2.5}$.

Solution

Using the arithmetic form:

$$\begin{aligned} 10^{2.5} &= 10^{\frac{5}{2}} = 10^{5\times\frac{1}{2}}(10^5)^{\frac{1}{2}} \\ &= \sqrt[2]{10^5} = \sqrt[2]{100{,}000} \\ &= 316.227 \end{aligned}$$

Using the exponential form:

$$\begin{aligned} 10^{2.5} &= 10^{\frac{5}{2}} = 10^{5\times\frac{1}{2}} = \left(10^{\frac{1}{2}}\right)^5 = (\sqrt[2]{10})^5 \\ &= \sqrt{10}\times\sqrt{10}\times\sqrt{10}\times\sqrt{10}\times\sqrt{10} \\ &= 316.227 \end{aligned}$$

Using the logarithmic form:

$$\begin{aligned}10^{2.5} &= 10^{2+0.5} = 10^{2}\times 10^{0.5}\\ &= 10^{2}\times\sqrt{10} = 10^{2}\times\sqrt[2]{10}\\ &= 100\times 3.16227\\ &= 316.227\end{aligned}$$

This example shows that complex exponents may be subdivided into powers and roots or, where a common denominator can be found, into a series of roots. If we are comfortable with addition, subtraction, multiplication, division, powers, and roots, we have an adequate knowledge of arithmetic so far as audio systems are concerned. If, for example, we find that the addition of negative numbers causes us difficulty, then some review is well worthwhile.

6.8 Complex Number Operations

To work with complex numbers, refer to Fig. 6-3.

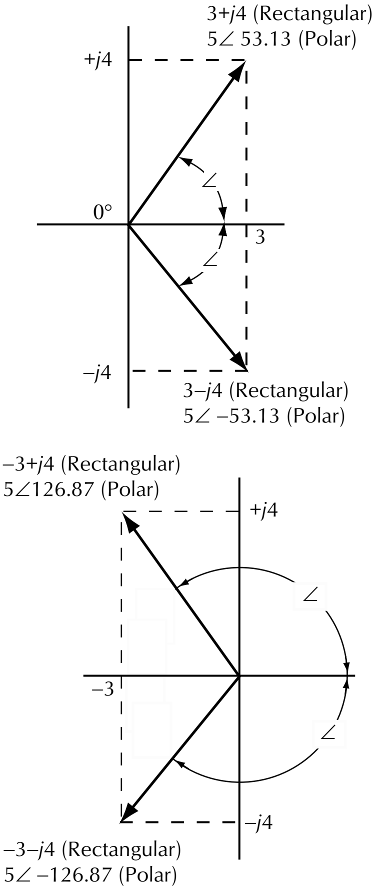


Figure 6-3. Complex numbers expressed in rectangular form.

1. To add complex numbers, use rectangular form ($a + jb$). Add real parts and then add imaginary parts.

$$\begin{array}{r} 3+j4\\ \underline{+3-j4}\\ 6+j0\\ =6 \end{array}$$

2. To subtract complex numbers, use rectangular form. Subtract real parts and then subtract imaginary parts.

$$\begin{array}{r} 3+j4\\ \underline{-(3-j4)}\\ 0+j8\\ =j8 \end{array}$$

3. To multiply complex numbers, use the polar form, ($M\angle\theta$ or $Me^{j\theta}$), where $M = (a^2 + b^2)^{1/2}$. Multiply the magnitudes and add the angles. To calculate magnitude and angle, use Eqs. 6-25 and 6-26.

$$\begin{array}{rcr} (3+j4)(3-j4) & = & 5\angle 53.13\\ & & \underline{5\angle -53.13}\\ & & 25\angle 0(0°) = 25 \end{array}$$

4. To divide complex numbers, use the polar form, ($M\angle\theta$). Divide magnitudes and subtract angles.

$$\begin{aligned}\frac{3+j4}{3-j4} &= \frac{5\angle 53.13}{5\angle -53.13}\\ &= 1\angle 106.26\\ &= (-0.28+j0.96)\end{aligned}$$

5. To obtain a complex number raised to the power n use the polar form. Raise the magnitude to the power n, (M^n), and then multiply the angle by n, ($e^{in\theta}$).
6. To obtain n roots of complex numbers, use the polar form ($M\angle\theta$) or $Me^{j\theta}$. Extract the root of the magnitude ($M^{1/n}$). Divide the angles by n. Example: To find the cube root of 8∠90°, $8^{1/3} = 2$. Then 90°/3 = 30°, (90° + 360°)/3 = 150°, and (90° = 720°/3 = 270°). The three roots are then 2∠30°, 2∠150°, and 2∠270°.

7. We can use the foregoing method to calculate impedance. If we measured an ac resistance R of 12.26 Ω and a total reactance X of 10.28 Ω we could plot it as shown in Fig. 6-3. The angle θ is called the angle of impedance and, in this illustration, is the number of degrees or radians the voltage leads the current. ELI THE ICE MAN is an easy way to remember that E (voltage) leads I (current) when L (inductance) is involved and that I (current) leads E (voltage) when C (capacitance) is involved. M from the previous

examples is now the magnitude of the impedance. We would write the data as:

$$R \pm jX = 12.26 + j10.28 \tag{6-25}$$

To find the magnitude of the impedance we can use the Pythagorean theorem:

$$Z = \sqrt{R^2 + X^2} \tag{6-26}$$

and trigonometrically:

$$\theta = \tan^{-1}\left(\frac{X}{R}\right) \tag{6-27}$$

($\tan^{-1}$ simply means the inverse of the tangent).

It is problems of this type that the advent of low cost scientific electronic calculators have made so accessible without undue stress and strain on the mental system.

6.9 Decade Calibration

A decade in history is 10 years, but a decade in audio is any ten-part interval (decade resistance boxes have controls calibrated from 0 to 9 on each knob). A decade in frequency would be defined as:

$$\frac{H.F.}{L.F.} = 10^1 = 1 \text{ decade} \tag{6-28}$$

where,
H.F. is the highest frequency,
L.F. is the lowest frequency.

This means that we can write a general case expression for the calculation of how many decades there are in a given bandpass by:

$$\frac{H.F.}{L.F.} = 10^{(x \text{ decades})} \tag{6-29}$$

and

$$\frac{\ln H.F. - \ln L.F.}{\ln 10} = x \text{ decades} \tag{6-30}$$

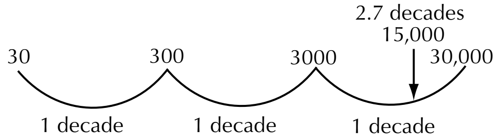
where,
ln is the natural logarithm or the logarithm to the base *e*.

Example

Using a frequency span of 30–15,000 Hz in discussing octaves, we can see that it represents 2.7 decades:

$$x \text{ decades} = \frac{\ln 15{,}000 - \ln 30}{\ln 10}$$

$$= 2.7 \text{ decades}$$



The question might arise in a different form. For example, what upper frequency limit would I have if I extended 2.5 decades from 30 Hz?

$$H.F. = e^{\ln 10^{(2.5)} + \ln 30} = 9486.8 \text{ Hz}$$

Or, what if I wish a 2.5 decade span with an upper limit of 5000 Hz? What is my low-frequency cutoff?

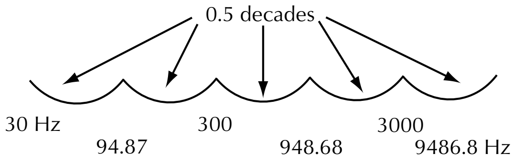
$$L.F. = e^{[\ln 5000 - \ln 10^{(2.5)}]} = 15.8 \text{ Hz}$$

Division into decades is electronically convenient (witness ¹⁄₁₀ decade rather than ⅓ octave filter designs), the general case equation for equally spaced logarithmic intervals can be derived from our approach above:

$$\left[\frac{H.F.}{L.F.}\right]^{\frac{1}{N}} = (multiplier) \tag{6-31}$$

where,
N is the number of intervals of equal logarithmic spacing (i.e., on a log scale are equally spaced).

Using the earlier data, suppose we want each frequency from 30 to 9486.8 Hz spaced ½ decade apart. That means we want $N = 5$.



Using equation 6-31:

$$\left[\frac{9486.8 \text{ Hz}}{30 \text{ Hz}}\right]^{\frac{1}{5}} = 3.16$$

As shown in Fig. 6-4, the first mark is 30 Hz; the second mark is 94.87, which is 30 × 3.16, while the third frequency is 300 or 3.16 × 94.87. The fourth frequency is 948.68 or 300 × 3.16, and the next frequency is 3000. The last frequency is 9486.8 Hz, where 2.5 are spaced at ½ decade calibration.

6.10 Converting Linear Scales to Logarithmic Scales

A linear frequency scale is one on which each equal length division represents an equal number of Hz. Thus, addition of this equal number of Hz to the last frequency gives the next frequency in the series.

A logarithmic scale is one on which equal length divisions represent an exponential constant. Thus, multiplying the last frequency by the exponential constant gives the next frequency in the series.

Example

To divide a linear scale into six equal parts (seven points including the first and the last with six equal length intervals enclosed between the points) with the maximum value equal to 300 and the beginning at zero, divide the maximum value by the number of equal intervals desired:

$$\frac{300}{6} = 50/\text{interval}$$

The first point is 0; the second point is 50 (enclosing one interval); the third point is 100 (enclosing the second interval); the fourth point is 150 (enclosing the third interval); the fifth point is 200 (enclosing the fourth interval); the sixth point is 250 (enclosing the fifth interval); and the seventh point is 300 (enclosing the sixth interval).

Example

To divide the same length scale into six equal logarithmic intervals from 1 to 300, find the multiplier constant:

$$\left[\frac{300}{1}\right]^{\frac{1}{6}} = 2.59$$

The first point is then 1.0; the second point is 2.59 (enclosing the first interval); the third point is 6.69 (2.59 × 2.59) (enclosing the second interval); the fourth point is 17.32 (2.59 × 2.59 × 2.59) (enclosing the third interval); the fifth point is 44.8 (2.59^4) (enclosing the fourth interval); the sixth point is 115.95 (2.59^5) (enclosing the fifth interval); the seventh point is 300 (2.59^6) (enclosing the sixth interval). These two scales are illustrated in Fig. 6-5.

Both linear and logarithmic scales are used repeatedly in sound system engineering. One-third octave scales are logarithmic spaced intervals. Signal delay anomalies are equally spaced on linear scales. The decibel scale is a logarithmic scale. The voltmeter scale is a linear scale. We need to cultivate familiarity with both types of scales and often real insights are gained by transferring data from one of these scales to the other.

6.11 Finding the Renard Series for Fractional Octave Spacing

Renard numbers are equally spaced intervals on a logarithmic scale. In the number system, one octave

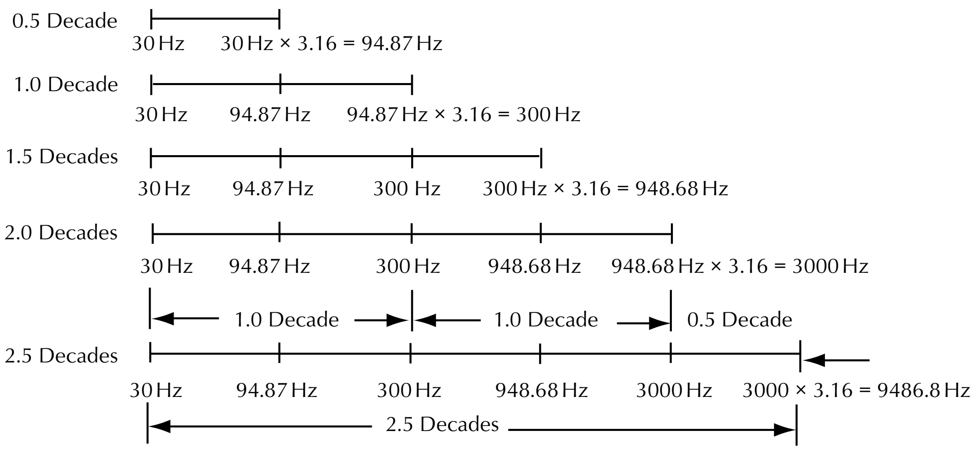


Figure 6-4. Decade calibration.

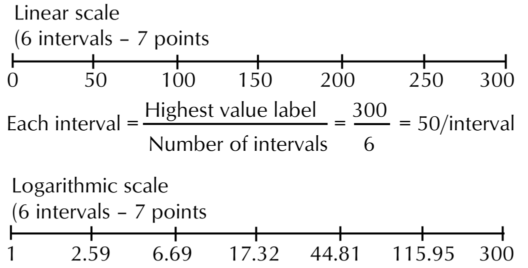


Figure 6-5. Linear and logarithmic scaling of equally spaced intervals from 0 to 300.

is specified as the 3⅓ series (i.e., $10^{1/3.33}$ is the multiplier m used to increment each interval). To find any other Renard series, multiply 3⅓ by the reciprocal of the fractional octave desired: (i.e.,
⅙ octave = 6 × 3⅓ = 20 series;
⅒ octave = 10 × 3⅓ = 33⅓ series):

$$f \text{ intervals} = (L_{f \times m})m \ldots m \tag{6-32}$$

Example

Obtain ⅓ octave intervals from 100 Hz up.

Solution

$$f \; intervals = 100\left[\left(10^{\frac{1}{10}}\right) \times \left(10^{\frac{1}{10}}\right) \times \left(10^{\frac{1}{10}}\right) \times \ldots \left(10^{\frac{1}{10}}\right)\right]$$

for a total of n multiplications where n is the number of intervals

6.11.1 Determining the Number of Octaves in a Given Bandwidth

One octave is a 2 to 1 change in frequency such that from 1 Hz to 2 Hz is one octave. So is 1000 Hz to 2000 Hz or 10,000 Hz to 20,000 Hz. This can be written mathematically as:

$$\frac{f_H}{f_L} = 2^n \tag{6-33}$$

where,
f_H is the higher frequency,
f_L is the lower frequency,
n is the number of octaves.

For example, 4 Hz/1 Hz = 2^2 or 2 octaves. Using natural logarithms, we can easily formulate this relationship into a general case equation:

$$\ln\left[\frac{f_H}{f_L}\right] = \ln 2 \times n \tag{6-34}$$

therefore,

$$n = \frac{\ln\left[\frac{f_H}{f_L}\right]}{\ln 2} \tag{6-35}$$

Example

How many octaves are there between 30 Hz and 15,000 Hz?

Solution

Using Eq. 6- 35,

$$n = \frac{\ln\left[\frac{15,000}{30}\right]}{\ln 2} = 8.97 \text{ octaves}$$

This is a very wide range high fidelity system (measured not advertised). Very few technological systems are asked to span such a range. All of visible light is but one octave. Audio is more complex than is often realized.

6.12 Radians and Steradians

6.12.1 The Radian

The radian is the plain angle between two radii of a circle that cuts off, on the circumference, an arc equal to the radius from the center of that circle to the circumference.

As a full circle is 2π radians which = 360°, then
π is a half circle,
π/2 is a 1/4 circle,
π/3 is a 1/6 circle,
etc. See Fig. 6-6.

Fig. 6-7 is the MathCAD program printout of radians and degrees.

6.12.2 Steradians (Solid Angles)

The "solid angle" is the angle that, seen from the center of a sphere, includes a given area on the surface of that sphere. While dimensionless the solid angle is given a label steradians (sr). One steradian is

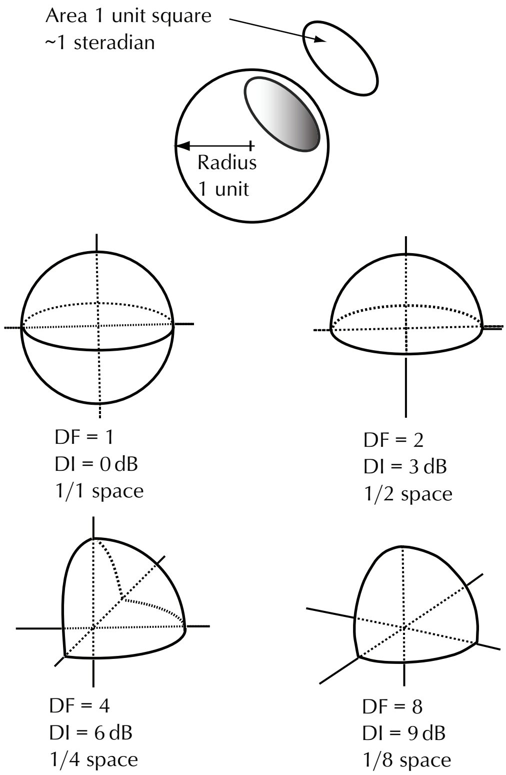


Figure 6-6. Directivity factors and directivity indexes.

the solid angle that cuts off an area of the surface of a sphere equal to that of a square with sides of length equal to the radius of the sphere when its vertex is in the center of the sphere. The numerical value of the solid angle is equal to the size of that area divided by the square of the radius of the sphere.

6.12.3 Symbols

1. *sr* is the solid angle in steradians.
2. *A* is the area on the surface of the sphere.
3. *r* is the radius of the sphere.
4. *rd* is the angle (∠) in radians.
5. *Q* is the geometric directivity factor.
6. $C_\angle$ is the equivalent coverage angle for squares or rectangles on the sphere's surface (also expressed as θ for horizontal angles and ϕ for vertical angles).

Caution

When working with trigonometric functions, computers and calculators need to select either the radians or degrees mode. Attention must be paid to the use of π with radians and 180° with degrees in the equations where trigonometric functions occur.

6.12.4 Equations

$$A = (sr)r^2 = \frac{4\pi r^2}{Q} \tag{6-36}$$

$$sr = \frac{A}{r^2} = \frac{4\pi}{Q} \tag{6-37}$$

$$r = \sqrt{\frac{A}{sr}} = \sqrt{\frac{QA}{4\pi}} \tag{6-38}$$

$$Q = \frac{4\pi}{sr} = \frac{4\pi r^2}{A} \tag{6-39}$$

$$Q = \frac{\pi}{\operatorname{asin}\left[\sin\left(\frac{\theta}{2}\right) \times \sin\left(\frac{\phi}{2}\right)\right]} \tag{6-40}$$

where θ and ϕ are in radians

$$Q = \frac{180}{\operatorname{asin}\left[\sin\left(\frac{\theta}{2}\right) \times \sin\left(\frac{\phi}{2}\right)\right]} \tag{6-41}$$

where θ and ϕ are in degrees

$$sr = 4 \times \arcsin\left[\frac{\sin\theta}{2} \cdot \frac{\sin\phi}{2}\right] \tag{6-42}$$

where θ and ϕ are in radians

$C_\angle\theta$ and ϕ from a *Q* for a square area

$$C_\angle = 2\operatorname{asin}\sqrt{\sin\left(\frac{\pi}{Q}\right)} \tag{6-43}$$

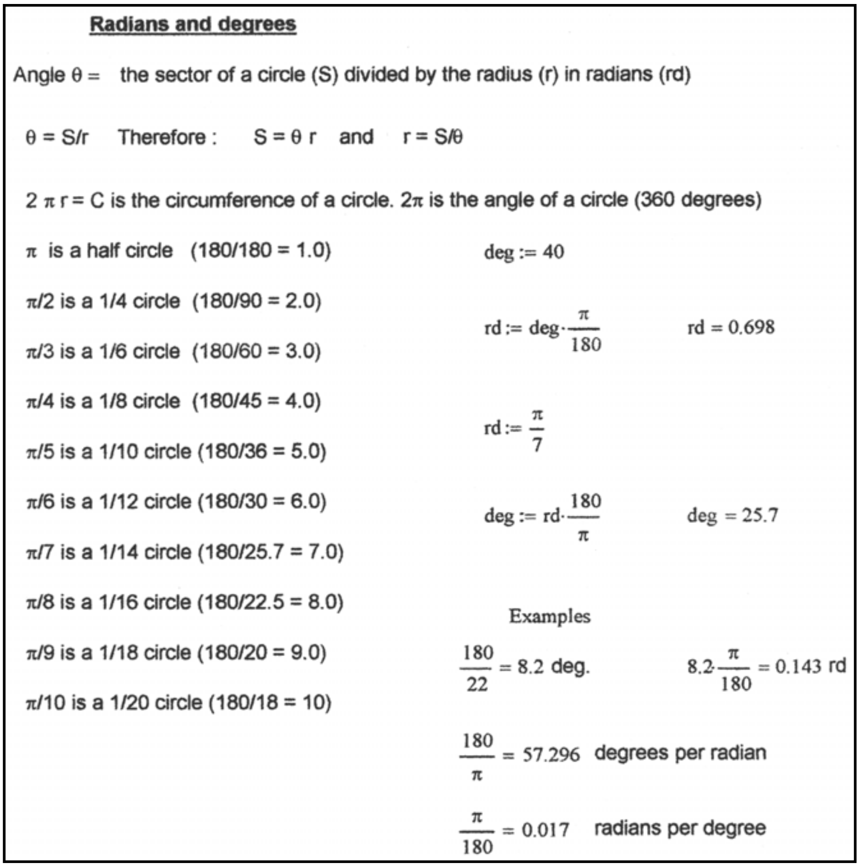
radian mode

$$C_\angle = 2\operatorname{asin}\sqrt{\sin\left(\frac{180}{Q}\right)} \tag{6-44}$$

degree mode

6.12.5 Complementary Angle for a Given Square Angle and a Desired Arbitrary Angle

Given:
$C_{\angle S}$ (square angle) in rd,



Radians and degrees

Angle θ = the sector of a circle (S) divided by the radius (r) in radians (rd)

θ = S/r Therefore : S = θ r and r = S/θ

2 π r = C is the circumference of a circle. 2π is the angle of a circle (360 degrees)

π is a half circle (180/180 = 1.0)

π/2 is a 1/4 circle (180/90 = 2.0)

π/3 is a 1/6 circle (180/60 = 3.0)

π/4 is a 1/8 circle (180/45 = 4.0)

π/5 is a 1/10 circle (180/36 = 5.0)

π/6 is a 1/12 circle (180/30 = 6.0)

π/7 is a 1/14 circle (180/25.7 = 7.0)

π/8 is a 1/16 circle (180/22.5 = 8.0)

π/9 is a 1/18 circle (180/20 = 9.0)

π/10 is a 1/20 circle (180/18 = 10)

$\text{deg} := 40$

$\text{rd} := \text{deg}\cdot\frac{\pi}{180}$ $\text{rd} = 0.698$

$\text{rd} := \frac{\pi}{7}$

$\text{deg} := \text{rd}\cdot\frac{180}{\pi}$ $\text{deg} = 25.7$

Examples

$\frac{180}{22} = 8.2$ deg. $8.2\cdot\frac{\pi}{180} = 0.143$ rd

$\frac{180}{\pi} = 57.296$ degrees per radian

$\frac{\pi}{180} = 0.017$ radians per degree

Figure 6-7. Radians and degrees.

$C_{\angle A}$ (arbitrary angle) in rd,
$C_{\angle C}$ (complementary angle) in rd.

$$C_{\angle C} = 2\,\text{asin}\left[\frac{\sin\left(\frac{C_{\angle S}}{2}\right)^2}{\sin\left(\frac{C_{\angle A}}{2}\right)}\right] \text{ in radians}$$

Conversions (6-45)

$$\text{rd}\cdot\frac{180}{\pi} = degrees$$

$$degrees\cdot\frac{\pi}{180} = \text{rd}$$

6.12.6 Finding Q for a Given Square Angle

$$Q = \frac{\pi}{\text{asin}\left[\sin\left(\frac{C_{\angle S}}{2}\right)^2\right]} \tag{6-46}$$

rd mode

$$Q = \frac{180}{\text{asin}\left[\sin\left(\frac{C_{\angle S}}{2}\right)^2\right]} \tag{6-47}$$

degree mode

6.12.7 Percentage of Spherical Surface Area Covered for a Given Q

$$\begin{aligned}\%\text{ Area} &= 100\left(\frac{1}{Q}\right)\\ &= 100\left(\frac{sr}{4\pi}\right)\end{aligned} \tag{6-48}$$

$$\begin{aligned}1.0sr &= \text{square angle of } 1.041198 \text{ rd}\\ &= 59.65589°\\ &= Q \text{ of } 4\pi\\ &= 7.958\% \text{ of spherical surface area}\end{aligned}$$

6.13 Calculating Percentages and Ratios

The story is told of the “least likely to succeed” member of a class returning to the class reunion. He arrived in a Rolls Royce attended by a chauffeur, footman, bodyguards, etc. Someone asked him how he made so much money. He replied, “I found a small item everyone wanted and marked it up 10% and the money rolled in. I bought it for a dollar and sold it for $10.00, and that 10% really added up.”

Fig. 6-8 and the following equations are used in financial percentage problems:

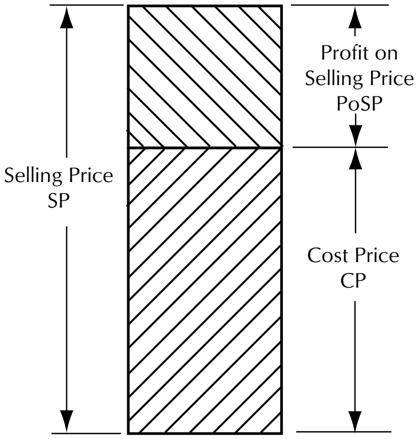


Figure 6-8. Financial percentage terms.

$$\%PoSP = 100\left(1 - \frac{CP}{SP}\right) \tag{6-49}$$

$$CP = SP\left(1 - \frac{\%PoSP}{100}\right) \tag{6-50}$$

$$CP = \frac{SP}{MR^*} \tag{6-51}$$

$$SP = \frac{CP}{1 - \frac{\%PoSP}{100}} \tag{6-52}$$

$$SP = MR^* \times CP \tag{6-53}$$

$$MR^* = 100\left(\frac{1}{100 - \%PoSP}\right) \tag{6-54}$$

$$MR^* = \frac{SP}{CP} \tag{6-55}$$

* Markup ratio

$$\%MoCP = 100\left(\frac{SP}{CP} - 1\right) \tag{6-56}$$

where,
%PoSP is the % profit on selling price,
%MoCP is the % markup on the cost price,
SP is the selling price,
CP is the cost price,
MR is the markup ratio.

Example

Suppose you pay $50.00 for an item (*CP*) and sell it for $75.00 (*SP*). What is your % profit on selling price (*%PoSP*) on that sale?

Solution

Using Eq. 6-49:

$$\%PoSP = 100\left(1 - \frac{50}{75}\right)$$
$$= 33\tfrac{1}{3}$$

Example

Suppose you are buying an article for $75.00 and later are told that the dealer makes 33⅓% on it. What did the dealer pay for the article (*CP*)?

Solution

Using equation 6-50:

$$CP = \$75.00\left(\frac{100\% - 33.333\%}{100\%}\right)$$
$$= \$50.00$$

Example

Suppose you pay $50.00 for an article and wish to have a 33⅓% profit. What selling price should you receive to do this?

Solution

Using Eq. 6-52:

$$SP = \frac{\$50.00}{1 - \frac{33.33\%}{100}}$$
$$= \$75.00$$

Note here that it is not possible to make a 100% profit on the selling price unless you receive the article as a gift. These questions can also arrive as markup ratios *MR* or as percent markup on cost price *%MoCP*.

Example

Suppose you pay $50.00 for the article, sell it for $75.00, and realize a 33⅓% profit. What is the markup ratio?

Solution

Using Eq. 6-54:

$$MR = 100\left(\frac{1}{100 - 33.33}\right)$$
$$= 1.5$$

In other words, 1.5 × $50.00 = $75.00. Thus, any cost multiplied by 1.5 will give the correct selling price whenever the percent profit on selling price is to be 33⅓%.

Example

When the cost price is $50.00 and the selling price $75.00, what is the percent markup on cost price?

Solution

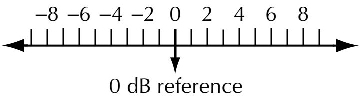
Using Eq. 6-56:

$$\%MoCP = 100\left(\frac{75.00}{50.00} - 1\right)$$
$$= 50\%$$

Using these equations, we can see that the successful class member had a 90% *PoSP* or a 900% *MoCP*, not a 10% profit as he had assumed.

6.13.1 Decibels and Percentages

The decibel indicates a change in ratio. Provided all voltages are measured across the same resistance or impedance voltage or current ratios can be found as well, Fig. 6-9.



Decibels above reference
$\%\text{ change} = 100(b^{+db/M}-1)$

Decibels below reference
$\%\text{ change} = 100(1-b^{-db/M})$

where,
b is the base (usually 10 or 2.718...),
M is the multiplier (usually 10, power, or 20, voltage.)

Figure 6-9. Scale with 0 dB reference.

6.13.2 Power Changes Downward

$$\%Change = 100\left(1 - 10^{\frac{-dB}{10}}\right) \tag{6-57}$$

6.13.3 Power Changes Upward

$$\%Change = 100\left(10^{\frac{+dB}{10}} - 1\right) \tag{6-58}$$

6.13.4 Voltage Changes Downward

$$\%Change = 100\left(1 - 10^{\frac{-dB}{20}}\right) \tag{6-59}$$

6.13.5 Voltage Changes Upward

$$\%Change = 100\left(10^{\frac{+dB}{20}} - 1\right) \tag{6-60}$$

Example

An anechoic chamber absorbs 99% of the power put into it and reflects only 1%. What percent of the initial sound pressure level (*SPL*) is reflected?

Solution

First, if we put 100 acoustic watts into the room you will have 1 W reflected:

$$10\log\frac{100}{1} = 20\text{ dB}$$

20 dB = 99% of the power absorbed.

Therefore, reflection will be 20 dB *below* where it started. Since the L_P is a voltagelike quantity, it also will be –20 dB:

$$\%\,Change = 100\left(1 - 10^{\frac{-20}{20}}\right) = 90\%$$

Thus 90% of the L_P was absorbed. Since we started with 100%:

$$100 - 90 = 10\%$$

which is the reflection percentage for L_P.

Example

Having equalized a sound system and raised its gain 15 dB, what percent power increase is now called for?

Solution

$$\%Change = 100\left(10^{\frac{+15}{10}} - 1\right) = 3062\%$$

6.13.6 Decibels for Percent Below Reference

$$dB = M\log_b\left(1 - \frac{\%Change}{100}\right) \tag{6-61}$$

6.13.7 Decibels for Percent Above Reference

$$dB = M\log_b\left(\frac{\%Change}{100} + 1\right) \tag{6-62}$$

Example

If a harmonic is 1% (i.e., 99% below reference)

$$\text{dB} = 20^{*}\log_{10}\left(1 - \frac{99\%}{100}\right) = 40\text{ dB}$$

* Voltagelike ratios

Example

An acoustic signal is reflected off of a surface that is 80% absorptive, the reflected signal will drop 6.99 dB.

$$\text{dB} = 10^{*}\log_{10}\left(1 - \frac{80}{100}\right) = 6.99\text{ dB}$$

* Powerlike ratios

Example

If the input voltage to loudspeaker is raised by 30%, we should add 2.28 dB to its L_P

$$\text{dB} = 20\log_{10}\left(\frac{30\%}{100} + 1\right) = +2.28\text{ dB}$$

Fig. 6-10 shows the relationship between decibels and percentages for power and for voltage, current, L_P, and distances.

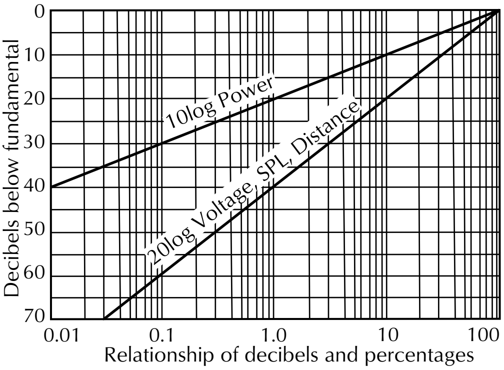


Figure 6-10. Relationship of decibels and percentages.

6.14 Useful Math Tables

A table of the basic relationships is shown in Table 6-6. Table 6-7 shows widely used exponential prefixes. Basic trigonometry is frequently useful in sound system work, Table 6-8. Table 6-9 illustrates a simple but practical application.

6.14.1 Scientific Metrology

It appears that a slow shift (beginning with active minorities such as cosmologist and string theorists) is underway in physics which will gradually cause the metric system to be abandoned in favor of practical versions of the Planck units. These will be conveniently scaled by powers of 10, the way scientific units always are, and some will be ascribed conventional values to provide exact metric convertibility.

Units

In a totally fascinating article in *Physics Today*, February 2003, by Frank Wilczek, the Herman Feshback Professor of Physics at MIT, entitled, “Life Parameters” is the best discussion of how to scale measurement parameters that I have ever read. I have a definite preference for “scaling” to human dimensions, and a “foot” is human indeed. The professor prefers CGS (grams, centimeter and seconds) for the same reason, rather than reflexing to SI (kilograms, meters and seconds).

In the United States of America there remains to the present time a preference for pounds, feet, seconds and their derivatives. Engineering conveniences aside, any audio engineer talking to a layman (translate clients) who presents room data in SI alienates that client. Most of you have lived jointly with the US and the SI systems without harm while being non-bilingual, whereas the Europeans are bilingual, yet they often stare blankly at US dimensions. I guess you could call Americans bi-dimensional.

Planck’s Units

The focus of Professor Wilczek’s article is Planck’s units. These are, expressed in his preferred CGS format, the smallest mass, the shortest length, and the shortest time.

1. 10^{-6} g for mass.
2. 10^{-33} cm for length.
3. 10^{-44} s for time.

Table 6-6. Trigonometric Functions in Terms of Each Other

$\sin\theta$	$\sin\theta$	$\sqrt{1-\cos^2\theta}$	$\frac{\tan\theta}{\sqrt{1+\tan^2\theta}}$	$\frac{1}{\sqrt{1+\cot^2\theta}}$	$\frac{\sqrt{\sec^2\theta-1}}{\sec\theta}$	$\frac{1}{\csc\theta}$
$\cos\theta$	$\sqrt{1-\sin^2\theta}$	$\cos\theta$	$\frac{1}{1+\tan^2\theta}$	$\frac{\cot\theta}{\sqrt{1+\cot^2\theta}}$	$\frac{1}{\sec\theta}$	$\frac{\sqrt{\csc^2\theta-1}}{\csc\theta}$
$\tan\theta$	$\frac{\sin\theta}{\sqrt{1-\sin^2\theta}}$	$\frac{\sqrt{1-\cos^2\theta}}{\cos\theta}$	$\tan\theta$	$\frac{1}{\cot\theta}$	$\sqrt{\sec^2\theta-1}$	$\frac{1}{\sqrt{\csc^2\theta-1}}$
$\cot\theta$	$\frac{\sqrt{1-\sin^2\theta}}{\sin\theta}$	$\frac{\cos\theta}{\sqrt{1-\cos^2\theta}}$	$\frac{1}{\tan\theta}$	$\cot\theta$	$\frac{1}{\sqrt{\sec^2\theta-1}}$	$\sqrt{\csc^2\theta-1}$
$\sec\theta$	$\frac{1}{\sqrt{1-\sin^2\theta}}$	$\frac{1}{\cos\theta}$	$\sqrt{1+\tan^2\theta}$	$\frac{\sqrt{1+\cot^2\theta}}{\cot\theta}$	$\sec\theta$	$\frac{\csc\theta}{\sqrt{\csc^2\theta-1}}$
$\csc\theta$	$\frac{1}{\sin\theta}$	$\frac{1}{\sqrt{1-\cos^2\theta}}$	$\frac{\sqrt{1+\tan^2\theta}}{\tan\theta}$	$\sqrt{1+\cot^2\theta}$	$\frac{\sec\theta}{\sqrt{\sec^2\theta-1}}$	$\csc\theta$

Table 6-7. SI Prefixes

Number	Power Ten Is Raised To	Prefix	Symbol
	10^{24}	Yotta	Y
	10^{21}	Zetta	Z
	10^{18}	Exa	E
	10^{15}	Peta	P
1,000,000,000,000.	10^{12}	Tera	T

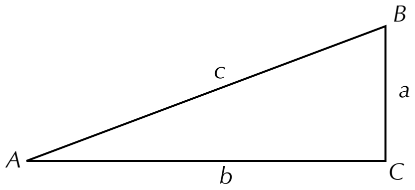
Table 6-7. (cont.) SI Prefixes

Number	Power Ten Is Raised To	Prefix	Symbol
1,000,000,000.	10^{9}	Giga	G
1,000,000.	10^{6}	Mega	M
1000.	10^{3}	kilo	k
100.	10^{2}	hecto	h
10.	10^{1}	deka	da

Table 6-7. (cont.) SI Prefixes

Number	Power Ten Is Raised To	Prefix	Symbol
1.	10^0	—	—
0.10	10^{-1}	deci	d
0.01	10^{-2}	centi	c
0.001	10^{-3}	milli	m
0.000 001	10^{-6}	micro-	µ
0.000 000 001	10^{-9}	nano	n
0.000 000 000 001	10^{-12}	pico	p
0.000 000 000 000 001	10^{-15}	femto	f
	10^{-18}	atto	a
	10^{-21}	zepto	z
	10^{-24}	yocto	y

Table 6-8. Basic Trigonometry



$$\sin A = \frac{a}{c} \qquad \csc A = \frac{c}{a}$$

$$\cos A = \frac{b}{c} \qquad \text{exsec}A = \sec A - 1$$

$$\tan A = \frac{a}{b} \qquad \text{vers}A = 1 - \cos A$$

$$\cot A = \frac{b}{a} \qquad \text{covers}A = 1 - \sin A$$

$$\sec A = \frac{c}{b} \qquad \text{hav}A = 0.5 versA$$

Radians	
$2\pi = 6.2832 = 360°$	$\pi/4 = 0.7854 = 45°$
$\pi = 3.1416 = 180°$	$\pi/6 = 0.5235 = 30°$
$\pi/2 = 1.5708 = 90°$	$\pi/12 = 0.26180 = 15°$
$\pi/3 = 1.0472 = 60°$	

The professor goes on in exquisite detail to relate mass to "brain mass" as limited by the mother's ability to give birth to a given size head. Length is worked on from the Bohr radius (a centimeter is roughly a 10^8 Bohr radii or atomic sizes), and so a cubic centimeter is just what encompasses those same 10^{24} atoms that make a gram. Time is immediately humanized by "So, why does it take about a second to have a thought?" Also how gravity relates to restricting the pace of human movement. Planck's units are constructed from suitable combinations of:

1. C = the speed of light.
2. h = the quantum of action.
3. G = the Newtonian gravitational constant.

These quantities are the avators of Lorentz symmetry: wave—particle duality, and the bending of space time by matter, respectively.

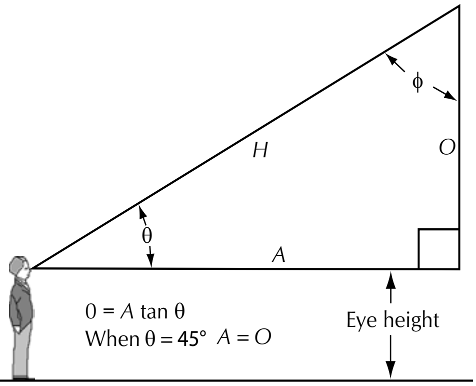
I am once again reminded that Wigner deplored that physics, or any other entity for that matter, had ever been able to define consciousness. When one works in a dimensional system of human proportions, wrong answers often appear ridiculous, whereas a system far removed from human sense perception often befuddles the user. The following reveals a richness that decimals will never attain:

8 furlongs = 1 mile = 5280 feet = 1760 yards = 8000 links = 320 rods = 80 chains = 0.86838249 nautical miles = 880 fathoms = 0.289 league = 63,360 inches.

The acre was originally a measure of a field's production rather than its dimension. All these dimensions were the basis of property division in the Western United States (the South preferred a more human-based system of "meets and bounds," i.e., from that tree to that corner, etc.) This has ingrained itself in the American psyche so solidly as to require draconian measures by a dictatorial power to change it.

My sense of all this is that SI is just one more set of dimensions to be aware of and when I want to be truly scientific, I'll resort to Planck's units.

Table 6-9. Finding an Unknown Height



Walk back from elevation to be measured until q is at 45° or, if that is not possible, at some angle between 15° and 45°. Adjust your distance *A* to allow a whole angle to be measured. Add your eye height to *O*

∠	tanθ	∠	tanθ	∠	tanθ
15°	0.2679	26°	0.4877	37°	0.7536
16°	0.2867	27°	0.5095	38°	0.7813

Table 6-9. (cont.) Finding an Unknown Height

17°	0.3057	28°	0.5317	39°	0.8098
18°	0.3249	29°	0.5543	40°	0.8391
19°	0.3443	30°	0.5774	41°	0.8693
20°	0.3640	31°	0.6009	42°	0.9004
21°	0.3839	32°	0.6249	43°	0.9325
22°	0.4040	33°	0.6494	44°	0.9657
23°	0.4245	34°	0.06745	45°	1.000
24°	0.4452	35°	0.7002		
25°	0.4663	36°	0.7265		

6.15 Angles

Fig. 6-11 represents a sector of a circle or "a piece of pie." It is also a piece of pi (π) as will become evident in the following.

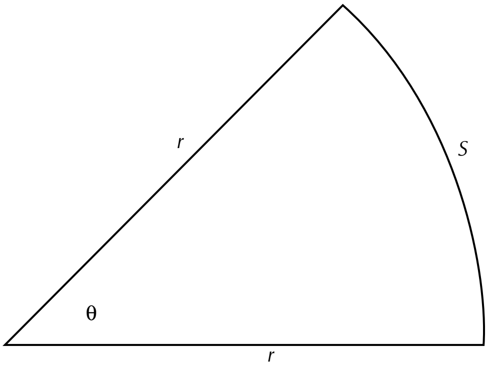


Figure 6-11. Sector of a circle.

The radius of the circle from which the sector was taken is r, the portion of the circle's perimeter belonging to the sector is the arc length S. The angle associated with the sector is θ and by definition,

$$\theta = \frac{S}{r} \tag{6-63}$$

It is apparent from the definition that θ is given by the ratio of two lengths and hence is itself dimensionless. Imagine now that one examines larger and larger sectors from the same circle. In fact, one could go "full circle" and examine the entire circle. In this instance S has become the perimeter of the circle, which is $2\pi r$ and θ has become

$$\theta = \frac{2\pi r}{r} = 2\pi \tag{6-64}$$

Going "full circle" yields an angle of 2π. One may well ask, "two pi what?" Even though θ is dimensionless, a nametag is usually appended to angular measure. This tag is the radian. Going "full circle" therefore involves an angle of 2π radians.

Imagine now that one starts with a very small sector with θ approximately zero and that θ is allowed to grow uniformly with the passage of time such that at any time t, θ is given by

$$\theta = \omega t \tag{6-65}$$

Here, ω is the uniform growth rate of θ with elapsed time, i.e.,

$$\omega = \frac{\theta}{t} \tag{6-66}$$

What time is required for θ to go "full circle"? This time is called the period T and from Eq. 6-66

$$T = \frac{2\pi}{\omega} \tag{6-67}$$

The reciprocal of the period is called the frequency, f, and represents the number of times in one second that θ goes "full circle."

$$\begin{aligned}\frac{1}{T} &= f \\ &= \frac{\omega}{2\pi}\end{aligned} \tag{6-68}$$

where,

f is measured in reciprocal seconds or Hertz,

ω is called the angular frequency or radian frequency because from Eq. 6-68

$$\omega = 2\pi f \tag{6-69}$$

where ω is expressed in radians per second.

Alternatively, one might begin with a sector for which θ has some initial value, say α, but in which θ is allowed to grow uniformly with time at the same rate as before. In this instance the value of θ will be given by

$$\theta = \omega t + \alpha \tag{6-70}$$

Relations (Eq. 6-67), (Eq. 6-68), and (Eq. 6-69) will be the same as before because α is a constant angle independent of time. α is called the epoch angle, that is to say the angle from which one starts.

6.16 A Little Trigonometry

Fig. 6-12 is constructed from Fig. 6-11 by the addition of a perpendicular so as to form a right triangle within the sector. The right triangle has a hypotenuse of length r, an opposite side of length b, and adjacent side of length a. We are indebted to Pythagoras for determining the relationship which exists between r, a, and b.

$$r^2 = a^2 + b^2 \tag{6-71}$$

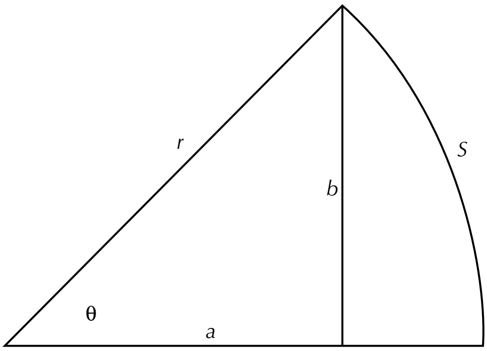


Figure 6-12. Sector containing a right triangle.

The sine of the angle θ written as $\sin\theta$ is defined to be

$$\sin\theta \equiv \frac{b}{r} \tag{6-72}$$

while the cosine of the angle θ is defined to be

$$\cos\theta \equiv \frac{a}{r} \tag{6-73}$$

and the tangent of the angle θ is defined to be

$$\tan\theta \equiv \frac{b}{a} \tag{6-74}$$

Even though the Eqs. 6-72, 6-73, and 6-74 are the defining equations for the $\sin\theta$, $\cos\theta$, and $\tan\theta$, respectively, there are other equivalent ways for determining the sine or cosine of a given angle. For example, the sine of the angle θ can be determined to any accuracy required by the following infinite series

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \ldots \tag{6-75}$$

while the cosine of the angle θ is given by the series

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \ldots \tag{6-76}$$

In Eqs. 6-75 and 6-76 the notation 2! means two multiplied by one, the notation 3! means three multiplied by two and then multiplied by one, and 4! means four multiplied by three then multiplied by two and finally multiplied by one, etc. The notation … means continued in the same manner ad infinitum.

Returning now to Eq. 6-72, it is observed that $\sin\theta$ as well as the other trigonometric functions are dimensionless. The utility of the equation exists in the fact that given r and θ one can determine b as

$$b = r\sin\theta \tag{6-77}$$

It is not necessary that r and b be distances as they are in the triangle of Fig. 6-12. They can have any dimensions you please as long as they have the same dimension and as long as they follow the triangular relationship for the right triangle given by Pythagoras.

6.17 The Origin of the Base of the Natural Logarithm, e

The statement that y is a function of x simply means that the value of a dependent variable which is y depends on the value of an independent variable which is x. Let us investigate one such functional relationship as expressed by the equation

$$y = (1 + x)^{\frac{1}{x}} \tag{6-78}$$

The equation states that in order to obtain y for a given x, you add one to x and raise the resulting quantity to the power, one divided by x. When x is very large, y is approximately one. When x is one, y is two. When x is 0.5, y is about 2.25. As x is allowed to get closer and closer to zero, y gets closer and closer to a number called e. In fact, e is the limiting value of y as x approaches zero in Eq. 6-78. This behavior is illustrated in Fig. 6-13.

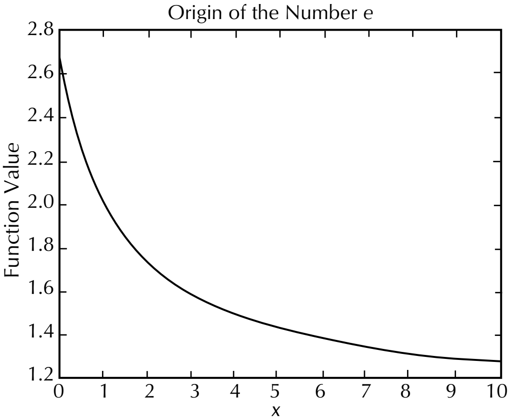


Figure 6-13. Base of the natural logarithm, e.

One of the consequences of defining e as the limiting value of Eq. 6-78 when x approaches zero is that e can be calculated to any degree of accuracy required from the infinite series

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \ldots \qquad (6\text{-}79)$$

Furthermore, e raised to any power such as u, is given to any degree of accuracy desired by

$$e^{u} = 1 + u + \frac{u^{2}}{2!} + \frac{u^{3}}{3!} + \frac{u^{4}}{4!} + \frac{u^{5}}{5!} + \ldots \qquad (6\text{-}80)$$

6.18 The Complex Plane

The normal algebraic operations such as addition, subtraction, multiplication, and division amount to locating points on the real number line of Fig. 6-14.

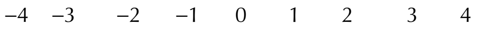


Figure 6-14. Real number line.

There is, however, no real number which when squared yields a negative real number. This is a consequence of how the operation of multiplication is defined for the numbers on the line of Fig. 6-14. We therefore invent a new kind of number whose square is a negative number. Historically such a number has been termed "imaginary" but really is no more imaginary than is any other number. This new type of number is distinguished by writing the symbol j in front of the number with the understanding that when j is multiplied by itself it yields -1, that is j is the square root of -1. A place is provided for these new numbers in a modification of Fig. 6-14 by extending a line upward and downward through zero thus arriving at a two dimensional space called the complex plane as depicted in Fig. 6-15.

Given a complex number having a real part a and an imaginary part b where the entire number is denoted by c, that is $c = a + jb$. What is c? c is a point in the complex plane which is located by beginning at the origin and then moving out a units on the real axis followed by then moving b units parallel to the imaginary or j axis. Alternatively, c could be located by moving r units along a line from the origin, where this line makes an angle θ with the real axis as shown in Fig. 6-16.

The relationships existing in Fig. 6-16 are as follows:

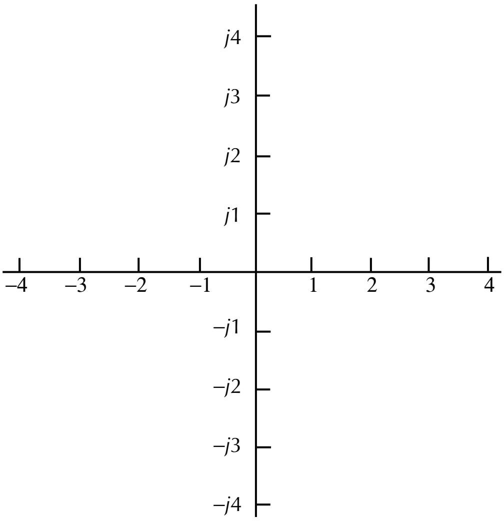


Figure 6-15. The complex plane.

1. $r = \sqrt{(a^{2} + b^{2})z}$, this is called the magnitude or modulus of the complex number c.
2. $\tan\theta = b/a$ where θ is called the angle of the complex number.

Note also that $a = r\cos\theta$ and $b = r\sin\theta$.

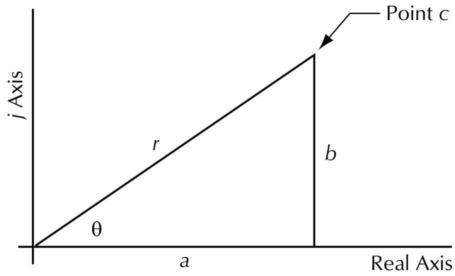


Figure 6-16. The complex quantity c.

6.19 Euler's Theorem

There is a famous theorem by Euler which states

$$e^{j\theta} = \cos\theta + j\sin\theta \qquad (6\text{-}81)$$

This can be easily proven. Take Eq. 6-80 and let $u = j\theta$ to obtain

$$e^{j\theta} = 1 + j\theta - \frac{\theta^{2}}{2!} - j\frac{\theta^{3}}{3!} + \frac{\theta^{4}}{4!} + j\frac{\theta^{5}}{5!} - \ldots \qquad (6\text{-}82)$$

Write Eq. 6-76

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \ldots$$

Multiply Eq. 6-75 by j to obtain

$$j\sin\theta = j\theta - j\frac{\theta^3}{3!} + j\frac{\theta^5}{5!} - j\frac{\theta^7}{7!} + \ldots \tag{6-83}$$

Now add Eq. 6-76 and 6-83 to yield

$$\cos\theta + j\sin\theta = 1 + j\theta - \frac{\theta^2}{2!} - j\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + j\frac{\theta^5}{5!} - \frac{\theta^6}{6!} - j\frac{\theta^7}{7!}\ldots \tag{6-84}$$

which is Eq. 6-82. Return now to the original complex number $c = a + jb$. Through the use of Euler's theorem, it should be apparent that c may be expressed in a variety of forms, each of which is equally valid

$$c = re^{j\theta} = r(\cos\theta + j\sin\theta) \tag{6-85}$$

$$c = a + jb = r\left(\frac{a}{r} + j\frac{b}{r}\right) \tag{6-86}$$

Eq. 6-86 is called the rectangular form for c whereas Eq. 6-85 is called the exponential form for c. They are equivalent in every respect. When adding or subtracting complex numbers, the rectangular form is the most convenient while the exponential form facilitates the operations of multiplication or division.

6.20 Examples

6.20.1 Addition of Complex Numbers.

Let $c = 5 + j7$ while $d = 4 + j3$. What is the sum $c + d$?

Step one: Add the real parts 5 + 4 to obtain 9.

Step two: Add the imaginary parts 7 + 3 to obtain 10.

Step three: Write $c + d = (5 + 4) + j(7 + 3) = 9 + j10$.

6.20.2 Subtraction of Complex Numbers.

Let $c = 5 + j7$ and $d = 4 + j3$. What is the difference $c - d$?

Step one: Subtract the real part of d from the real part of c to obtain $5 - 4 = 1$.

Step two: Subtract the imaginary part of d from the imaginary part of c to obtain $7 - 3 = 4$.

Step three: Write $c - d = (5 - 4) + j(7 - 3) = 1 + j4$.

Before proceeding to multiplication and division of complex numbers, c and d are converted to the exponential form.

$$\begin{aligned} c &= 5 + j7 \\ &= \sqrt{5^2 + 7^2}e^{j\,\mathrm{atan}\frac{7}{5}} \\ &= 8.6e^{j0.95} \end{aligned}$$

$$\begin{aligned} d &= 4 + j3 \\ &= \sqrt{4^2 + 3^2}e^{j\,\mathrm{atan}\frac{3}{4}} \\ &\ 5e^{j0.644} \end{aligned}$$

6.20.3 Products of Complex Numbers

The product of two complex numbers is a complex number whose magnitude is the product of the individual magnitudes and whose angle is the sum of the individual angles

$$cd = (8.6)(5)e^{j(0.95 + 0.644)} = 43e^{j(1.594)}$$

6.20.4 Quotients of Complex Numbers

The quotient of two complex numbers is a complex number whose magnitude is the quotient of the individual magnitudes and whose angle is the difference of the individual angles

$$\frac{c}{d} = \frac{8.6}{5}e^{j(0.95 - 0.644)} = 1.72e^{j(0.306)}$$

6.20.5 A Little Digression

Complex numbers are interesting and fun to play with and for most of their history were the province of pure mathematicians and later, a few physicists. An intellectual giant, Charles Proteus Steinmetz (1865–1923) was the first to introduce them into the engineering curriculum. Steinmetz arrived in the US in 1889 as a political refugee from Germany where he had been educated in mathematics, chemistry, and electrical engineering. He was hired as a consulting engineer by the fledgling General Electric Company and his contributions enabled GE to become an early industrial giant. Unfortunately, Steinmetz suffered from a congenital birth defect, which evidenced itself in the form of a hunched back. Steinmetz never married for fear of passing on his deformity. Nevertheless he loved young people and devoted many hours toward training young engineers as a professor at Union College all in addition to his work at GE. The introduction of the complex exponential notation into the electrical engineering curriculum greatly simplified ac circuit calculations and his mathematical techniques soon became standard industry wide. These techniques are not limited to just ac circuits but are equally applicable to any linear system in which the behavior has a sinusoidal time dependence.

6.21 Phasors

This section will be explored through an example dealing with the motion of a loudspeaker cone. Let x represent the displacement of the loudspeaker cone from its normal rest or equilibrium position. If a sinusoidal current of fixed frequency drives the loudspeaker, its displacement from equilibrium can be given by

$$x = x_m \cos(\omega t) \tag{6-87}$$

In this expression, x is considered to be positive when the cone moves so as to compress the air in front of it, x_m is the maximum value or amplitude of the displacement, and ω is the angular frequency. For the sake of definiteness, let

$$x_m = 10^{-6}\text{meter}$$

and

$$\omega = \frac{2\pi 1000}{s}$$

The loudspeaker displacement at any time is then

$$x = 10^{-6}\text{meter} \times \cos\left(\frac{2\pi 1000}{s}t\right)$$

This function can readily be represented by the complex exponential

$$x = 10^{-6}\text{meter} \times e^{j\left(\frac{2\pi 1000}{s}t\right)}$$

provided it is understood that only the real part of the complex exponential represents the true displacement. Recall, from Euler's theorem,

$$e^{j\theta} = \cos\theta + j\sin\theta$$

the real part of which is only cosθ. The function

$$x = 10^{-6}\text{meter} \times e^{j\left(\frac{2\pi 1000}{s}t\right)}$$

when depicted in the complex plane, is called a phasor. The depiction of this phasor at the instant $t = 0$ appears in Fig. 6-17A. The length or magnitude of this phasor is the same as the amplitude of the loudspeaker's cone motion. The angle of the phasor at any instant is

$$\theta = \frac{2\pi 1000}{s}t$$

so the phasor rotates counterclockwise about the origin at the rate

$$\omega = \frac{2\pi 1000}{s}$$

Fig. 6-17B displays the position of the phasor at an instant one-eighth of a period later than is displayed in Fig. 6-17A. Note that the projection of the phasor onto the real axis at any instant is

$$x = 10^{-6}\text{meter} \times \cos\left(\frac{2\pi 1000}{s}t\right)$$

which is the description of the loudspeaker motion.

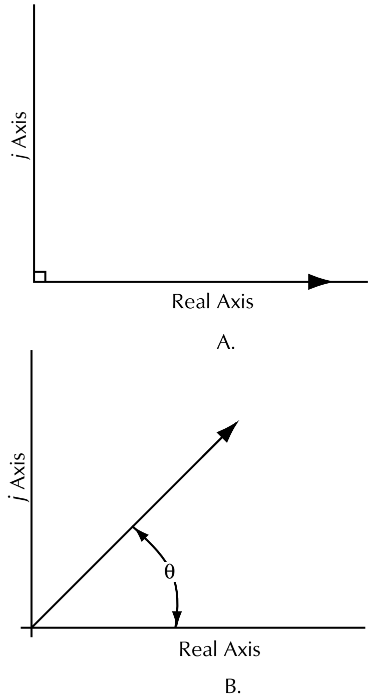


Figure 6-17. Phasor describing cone motion.

6.21.1 Addition of Phasors

Consider two loudspeakers operating at a frequency of 1000 Hz and positioned so that they each contribute to the total acoustic pressure at an observation point. The first loudspeaker acting alone produces an acoustic pressure given by the phasor

$$p_1 = 1 \text{ pascal } e^{j\left(2\pi\frac{1000}{s}t\right)}$$

The second loudspeaker acting alone produces an acoustic pressure given by the phasor

$$p_2 = 2 \text{ pascal } e^{j\left(2\pi\frac{1000}{s}t + \frac{\pi}{4}\right)}$$

Each of these acoustic pressures is small enough such that the air behaves linearly for these small disturbances. The total acoustic pressure at the observation point with both loudspeakers in operation will be the sum of the individual acoustic pressures. The total acoustic pressure at the observation point will be described by a phasor which is the sum of the individual phasors. Symbolically, the total acoustic pressure is described by

$$p_t = p_1 + p_2 \tag{6-88}$$

As each of the phasors is complex and complex quantities are most easily added when expressed in the rectangular form, the first step in finding the sum is to express each of the phasors in rectangular form. In doing this, it should be recalled that the length of a phasor does not depend upon the time of evaluation. Time does determine a phasor's angular position in the complex plane, but not its length. For convenience, the rectangular components are determined for the instant when $t = 0$. Therefore,

$$p_1 = 1 \text{ pascal } \cos(0) + j1 \text{ pascal } \sin(0)$$

$$p_2 = 2 \text{ pascal } \cos\left(\frac{\pi}{4}\right) + j2 \text{ pascal } \sin\left(\frac{\pi}{4}\right)$$

The next step is to add the real parts and separately to sum the imaginary parts. When the appropriate values for the sines and cosines are inserted, the result appears as

$$p_1 = 1 \text{ pascal} + j0$$

$$p_2 = \frac{2}{\sqrt{2}} \text{ pascal} + j\frac{2}{\sqrt{2}} \text{ pascal}$$

$$p_t = 2.4142 \text{ pascal} + j1.4142 \text{ pascal}$$

In the final step, the rectangular sum is converted to exponential form and the time dependence is re-inserted to obtain

$$p_t = 2.7979 \text{ pascal } e^{j\left(\frac{2\pi 1000}{s}t + 0.53\right)}$$

This process is depicted graphically in Fig. 6-18 which displays the individual phasors as well as their sum at the instant $t = 0$. As time increases, the entire figure rotates counterclockwise at the angular speed of 2000π radians per second. The actual acoustic pressure at the observation point is given by the real part of p_t and is

$$p_t = 2.7979 \text{ pascal } \cos\left(\frac{2\pi 1000}{s}t + 0.53\right)$$

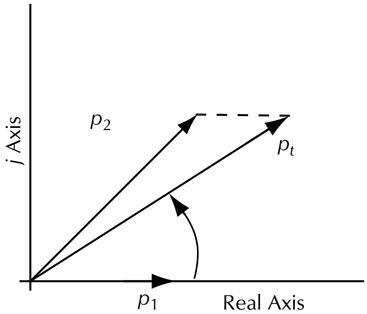


Figure 6-18. Phasor addition of acoustic pressures.

6.22 Rates of Change

Most of the laws of physics deal with the relationships between physical quantities and the rates at which these quantities change with respect to time or position, or both. As an example, Newton's second law of motion states that the time rate of change of linear momentum of a body is equal to the applied force. Linear momentum itself is the product of mass and linear velocity, but linear velocity is the time rate of change of linear position or displacement. In dealing with sinusoidal functions of time or position, or both, it is important to have mathematical tools for calculating the rates of change of these quantities. Lacking such a tool, Sir Isaac Newton was forced to develop one in the form of what is now known as differential calculus. The full power of the differential calculus will not be introduced here as it is possible to learn the answers at this point by employing only algebra and a little trigonometry in what follows. Attention should now be returned to the motion of the loudspeaker cone which was first described in the section on phasors

$$x = 10^{-6}\text{meter} \times \cos\left(\frac{2\pi 1000}{\text{s}}t\right)$$

Fig. 6-19 is a plot of the cone displacement versus time where the period has been divided into 20 equal increments of 5 × 10^{-5} s. This is the familiar cosine curve.

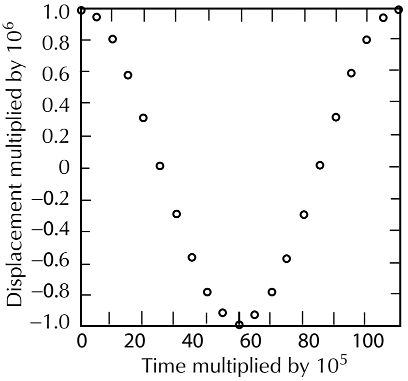


Figure 6-19. Loudspeaker cone displacement.

Fig. 6-20 is a curve constructed by taking successive differences between the points on Fig. 6-19 and dividing these differences by the time interval between points. Fig. 6-20 is thus a plot of change in position divided by change in time plotted versus time. This is known as the average velocity plotted versus time. Denote this average velocity as <*u*>.

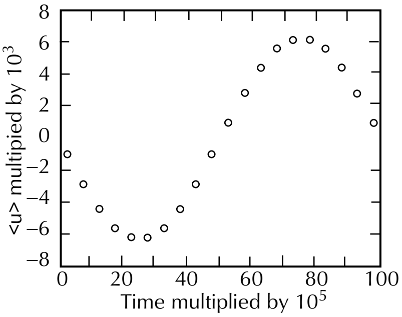


Figure 6-20. Average cone velocity.

The curve of <*u*> appears to be a negative sine curve. The information used in plotting Figs. 6-19 and 6-20 is summarized in Table 6-10. There are tabular entries for *t*, *x*, and <*u*> as calculated. There is an additional entry which represents

$$u = -2\pi 10^{-3}\,\frac{\text{meter}}{\text{s}}\sin\left(\frac{2\pi 10^3}{s}t\right)$$

If one performs the same analysis taking smaller and smaller time increments, it is observed that the discrepancy between columns 3 and 4 of Table 6-10 becomes smaller and smaller and in the limit column 3 becomes identical with column 4 which is the instantaneous velocity *u*.

The foregoing conclusion can be arrived at through a more analytical approach as follows. Use the phasor description of the motion of the loudspeaker cone.

$$x = x_m e^{j\omega t} \tag{6-89}$$

Multiply Eq. 6-89 by $j\omega$ to obtain

$$j\omega x = j\omega x_m e^{j\omega t} \tag{6-90}$$

Use Euler's identity to expand the complex exponential to obtain

$$j\omega x = \omega x_m (j\cos(\omega t) - \sin(\omega t)) \tag{6-91}$$

Finally, take only the real part of Eq. 6-91 and substitute numerical values to obtain

$$u = -2\pi 10^{-3}\,\frac{\text{meter}}{\text{s}}\sin\left(\frac{2\pi 10^3}{\text{s}}t\right) \tag{6-92}$$

Eq. 6-92 is the exact expression for the instantaneous velocity of the loudspeaker cone. The conclusion to be drawn is that the time rate of change of a sinusoidal quantity may be obtained by multiplying the phasor description of the quantity by $j\omega$. This process creates a new phasor, the real part of which is the true instantaneous time rate of change of the original sinusoidal time dependent quantity. In the language of the differential calculus, when working with phasors, the derivative with respect to time is obtained simply through multiplication by $j\omega$. In equation form, using calculus notation

$$u = \frac{dx}{dt} = j\omega x \tag{6-93}$$

Observe, Eq. 6-93 may be solved for *x* in terms of *u* to obtain

$$x = \frac{u}{j\omega}. \tag{6-94}$$

Table 6-10. Points of Figs. 1-19 and 1-20

Time (10^5)	x (10^6)	<u> (10^3)	u (10^3)
0	1		
2.5		−0.980	−0.982
5.0	0.951		

Table 6-10. (cont.) Points of Figs. 1-19 and 1-20

Time (10^5)	x (10^6)	$<u>$ (10^3)	u (10^3)
7.5		–2.840	–2.850
10.0	0.809		
12.5		–4.420	–4.440
15.0	0.588		
17.5		–5.580	–5.590
20.0	0.309		
22.5		–6.180	–6.200
25.0	0.000		
27.5		–6.180	–6.200
30.0	–0.309		
32.5		–5.580	–5.590
35.0	–0.588		
37.5		–4.420	–4.440
40.0	–0.809		
42.5		–2.840	–2.850
45.0	–0.951		
47.5		–0.980	–0.982
50.0	–1.000		
52.5		0.980	0.982
55.0	–0.951		
57.5		2.840	2.850
60.0	–0.809		
62.5		4.420	4.440
65.0	–0.588		
67.5		5.580	5.590
70.0	–0.309		
72.5		6.180	6.200
75.0	0.000		
77.5		6.180	6.200
80.0	0.304		
82.5		5.580	5.590
85.0	0.588		
87.5		4.420	4.440
90.0	0.809		
92.5		2.840	2.850
95.0	0.951		
97.5		0.980	0.982
100.0	1.000		

This implies that if the time rate of change of a quantity is known, the quantity itself can be obtained by dividing its time rate of change by $j\omega$. The process of undoing differentiation with respect to time is called integration with respect to time. This is a tool not of the differential calculus but rather of the integral calculus. In the language of the integral calculus

$$x = \int u dt = \frac{u}{j\omega} \tag{6-95}$$

This is the beauty of working with phasors. The mathematical operations of addition, subtraction, multiplication, division, differentiation, and integration can all be carried out employing only arithmetic, algebra, and a little trigonometry.

Bibliography

J. Backus. *The Acoustical Foundations of Music*. New York: Norton, 1969.

N. M. Cooke and H. F. R. Adams. *Basic Mathematics for Electronics*. New York: McGraw-Hill/Gregg Div., 1976.

A. H. Davis. *Modern Acoustics*. New York: Macmillan, 1934.

D. Davis. "Calculating Percentages and Ratios," *Syn-Aud-Con Tech Topics*, Vol. 6, No. 3 (1979).

O. W. Eschbach. *Handbook of Engineering Fundamentals*, 2nd ed. New York: Wiley, 1952.

W. Gellert, H. Kustner, M. Hellwich, and H. Kastner, Eds. *The VNR Concise Encyclopedia of Mathematics.* New York: Van Nostrand Reinhold, 1977.

H. L. F. Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, 2nd ed. New York: Dover, 1954.

ITT. *Reference Data for Radio Engineers*, 5th ed. Indianapolis, Ind.: Howard W. Sams, 1968.

M. Kaufman and A. H. Seidman, Eds. *Handbook of Electronics Calculations: For Engineers and Technicians.* New York: McGraw-Hill, 1979.

H. F. Olson. *Solutions of Engineering Problems by Dynamical Analogies*, 2nd ed. Princeton, N.J.: van Nostrand, 1943.

M. R. Schroeder. “Computers in Acoustics: Symbiosis of an Old Science and a New Tool,” *J. Acoust. Soc. Am.*, Vol. 45, No. 5 (1969).

————. *Number Theory in Science and Communication: With Applications in Cryptography, Physics, Biology, Digital Information, and Computing*. New York: Springer-Verlag, 1984.

R. Story. “Converting Tables to Equations Cuts Program Length for Calculator Use,” *Electronics*, Vol. 47 (Apr. 18, 1974), pp. 114-115.

S. P. Thompson, *Calculus Made Easy*. New York: Macmillan, 1937.

H. M. Tremaine, *Audio Cyclopedia*, 2nd ed. Indianapolis, Ind.: Howard W. Sams, 1969.

Chapter 7

Using the Decibel

by Don Davis

7.1 The Decibel 87
7.2 The Neper 87
7.3 Concepts Underlying the Decibel and Its Use in Sound Systems 88
  Converting Voltage Ratios to Power Ratios 88
  The dBV 88
  The Decibel as a Power Ratio 89
  The Decibel as a Power Quantity 90
7.4 Measuring Electrical Power 90
  Expressing Power as an Audio Level 90
7.5 Levels in dB 91
  Practical Variations of the dBm Equations 91
7.6 The Decibel in Acoustics—L_P, L_W, and L_I 92
7.7 Acoustic Intensity Level (L_I), Acoustic Power Level (L_W), and Acoustic Pressure Level (L_P) 93
  Acoustic Intensity Level, L_I 93
  Acoustic Power Level, L_W 93
  Acoustic Pressure Level, L_P 93
7.8 Inverse Square Law 93
7.9 Directivity Factor 94
7.10 Ohm's Law 94
7.11 A Decibel Is a Decibel Is a Decibel 95
  Older References 96
7.12 The Equivalent Level (*LEQ*) in Noise Measurements 96
7.13 Combining Decibels 97
  Adding Decibel Levels 97
  Subtracting Decibels 98
  Combining Levels of Uncorrelated Noise Signals 98
  To Add Levels 99
  To Subtract Levels 99
7.14 Combining Voltage 99
7.15 Using the Log Charts 99
  The 10Log *x* Chart 99
  The 20Log *x* Chart 100
7.16 Finding the Logarithm of a Number to Any Base 100
7.17 Semitone Intervals 101
7.18 System Gain Changes 101
7.19 The VU and the VI Instrument 101
  Crest Factor 102
  The VU Impedance Correction 102
  How to Read the VU Level on a VI Instrument 103
  Calibrating a VI Instrument 104
  Reading a VI Instrument on Program Material 104
  Reading Apparent VU Levels 104
7.20 Calculating the Number of Decades in a Frequency Span 104
7.21 Deflection of the Eardrum at Various Sound Levels 105
7.22 The Phon 105
7.23 The Tempered Scale 106
7.24 Measuring Distortion 106
7.25 The Acoustical Meaning of Harmonic Distortion 106

Calculating the Maximum Allowable Total Harmonic Distortion in an Arena Sound System 107
7.26 Playback Systems in Studios 108
Choosing an Amplifier 108
7.27 Decibels and Percentages 109
7.28 Summary 110
Bibliography 110

With a foundation in mathematics and physics, audio engineers can accurately describe sound and manipulate its effects with the decibel. The decibel, however, has worked its way into common usage. In the beginning, it was among several units of measurement.

Prior to 1923 "gains" and "losses" of telephone circuits were labeled in miles of standard cable (MSC). A "standard" cable was a 19 gauge open wire cable with a resistance of 88 Ω/mi and a capacitance of 0.054 µF⁄mi. The standard carbon transmitter's electrical output level with normal speech as an input was considered as "zero level" (at that time 0.006 W). The loss that occurred over one mile of this "standard" cable very nearly equaled what Harvey Fletcher was measuring in the Bell Telephone Laboratories as the smallest increment easily detected by a normal listener and which he labeled the sensation unit (*SU*).

In 1923 W. H. Martin wrote in the *Bell System Technical Journal* an article introducing the transmission unit (*TU*) devised to replace both the *MSC* and the *SU* while reconciling both uses to the telephone system as a whole. The *TU* was defined as:

$$(N)TU = 10\log\frac{P_1}{P_2} \qquad (7\text{-}1)$$

where,
TU was the transmission unit,
P_1 was the power measured,
P_2 was the power used as a reference,
N is the numerical value to be labeled as (*N*) *TU.*

There were 0.947 *MSC* to 1.0 *TU* or 1.056 *TU* to 1.0 *MSC.*

7.1 The Decibel

In 1929 W. H. Martin wrote again in the *Bell System Technical Journal* an article entitled "Decibel The Name for the Transmission Unit." As a result of Bell Telephone's participation as invited attendees, the European International Advisory Committee recommended to the various European telephone administrations that they adapt *either* the decibel or the Naperian unit and designate them the "Bel" and the "neper," respectively. The Bell system adopted the Bel, converting it into the deci-Bel (one tenth of a Bel) for the convenience of higher resolution of the measured level. Over the 20 years in which they had used the *MSC* they found that they could meaningfully resolve 0.1 mi. The decibel was defined as the logarithmic form of that power ratio having a value of $10^{0.1}$. Two amounts of power differ by 1 dB where they are in the ratio of $10^{0.1}$ and any two amounts of power differ by (*N*) dB when they are in the ratio of $10^{N(0.1)}$:

$$\begin{aligned}\frac{P_1}{P_2} &= 10^{N(0.1)}\\ \log_{10}\left(\frac{P_1}{P_2}\right) &= \log_{10}(10\times N\times 0.1)\\ \log_{10}10 &= 1\\ \frac{\log_{10}\frac{P_1}{P_2}}{0.1} &= N\\ 10\log_{10}\frac{P_1}{P_2} &= N\text{ dB}\end{aligned} \qquad (7\text{-}2)$$

where,
N units are labeled dB.

Hereafter it is assumed that log is to the base 10 in this book unless otherwise stated.

7.2 The Neper

The neper (*Np*) is defined as:

$$\ln\left(\frac{E_1}{E_2}\right) = N(\text{nepers}) \qquad (7\text{-}3)$$

One neper is a voltage *ratio* such that:

$$\begin{aligned}\left(\frac{E_1}{E_2}\right) &= e^1\\ &= 2.718\end{aligned}$$

N nepers is a voltage ratio such that:

$$\begin{aligned}\frac{E_1}{E_2} &= e^N\\ &= 2.718\ldots^N\end{aligned}$$

or alternatively:

$$\begin{aligned}\ln\left(\frac{E_1}{E_2}\right) &= \ln e\times N\\ \ln e &= 1\\ \ln\frac{E_1}{E_2} &= N\text{ (nepers)}\end{aligned}$$

The question arises as to what is the relationship between the neper and the decibel. Take the voltage ratio corresponding to one neper and square it to put it into the form of a power ratio and manipulate as follows:

$$\frac{P_1}{P_2} = \left(\frac{E_1}{E_2}\right)^2 \tag{7-4}$$

$$= e^2$$

$$10\log\frac{P_1}{P_2} = 10\log e^2$$

$$= \frac{8.686\text{ dB}}{\text{neper}}$$

The conclusion is that one neper is equivalent to 1/8.686 dB which is 0.115 Np/dB. Alternately you can use the decineper dNp (1 dB = 1.15 dNp).

7.3 Concepts Underlying the Decibel and Its Use in Sound Systems

Most system measurements of level start with a voltage amplitude. Relative level changes at a given point can be observed on a voltmeter scale when it is realized that:

$$10\log\frac{E_1^2}{E_2^2} = 10\log\frac{P_1}{P_2} \tag{7-5}$$

which is only true if both values are measured at an identical point in their circuit. A common usage has been to remove the exponent from the ratio and apply it to the multiplier.

$$2 \times 10\log\frac{E_1}{E_2} = 20\log\frac{E_1}{E_2} \tag{7-6}$$

Bear in mind that the decibel is always and only based upon a power ratio. Any other kind of ratio (i.e., voltage, current, or sound pressure) must first be turned into a power ratio by squaring and then converted into a power level in decibels.

7.3.1 Converting Voltage Ratios to Power Ratios

Many audio technicians are confused by the fact that doubling the voltage results in a 6 dB increase while doubling the power only results in a 3 dB increase. Fig. 7-1 demonstrates what happens if we simultaneously check both the voltage and power in a circuit where we double the voltage. Note that for a doubling of the voltage, the power increases four times.

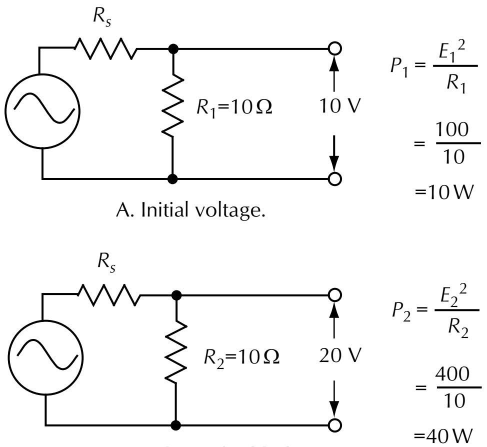


Figure 7-1. Voltage and power relationships in a circuit.

$$10\log\frac{P_1}{P_2} = 10\log\frac{40\text{ W}}{10\text{ W}}$$

$$= 6.02\text{ dB}$$

$$10\log\frac{E_1^2}{E_2^2} = 20\log\frac{20\text{ V}}{10\text{ V}}$$

$$= 6.02\text{ dB}$$

7.3.2 The dBV

One of the most common errors when using the decibel is to regard it as a voltage ratio (i.e., so many decibels above or below a reference voltage). To compound the error, the result is referred to as a "level." The word "level" is reserved for power; an increase in the voltage magnitude is properly referred to as "amplification."

However, the decibel can be legitimately used with a voltage reference. The reference is 1.0 V. When voltage magnitudes are referenced to it logarithmically, they are called dBV (i.e., dB above or below 1.0 V). This use is legitimate because *all* such measurements are made open circuit and can easily be converted into power levels at any impedance interface.

The following definition is from the *IEEE Standard Dictionary of Electrical and Electronics Terms, Second Edition*:

> **244.62**
>
> Voltage Amplification (1) (general). An increase in signal voltage magnitude in transmission from one point to another or the process thereof. See also: amplifier.

210 (2) (transducer). The scalar ratio of the signal output voltage to the signal input voltage. Warning: By incorrect extension of the term decibel, this ratio is sometimes expressed in decibels by multiplying its common logarithm by 20. It may be currently expressed in decilogs. Note: If the input and/or output power consist of more than one component, such as multi-frequency signal or noise, then the particular components used and their weighting must be specified. See also: Transducer.

239.210

Decilog (dg). A division of the logarithmic scale used for measuring the logarithm of the ratio of two values of any quantity. Note: Its value is such that the number of decilogs is equal to 10 times the logarithm to the base 10 of the ratio. One decilog therefore corresponds to a ratio of $10^{0.1}$ (that is 1.25829+).

7.3.3 The Decibel as a Power Ratio

Note that 20 W/10 W and 200 W/100 W both equal 3.01 dB, which means that a 2 to 1 (2:1) power ratio exists but reveals nothing about the actual powers. The human ear hears the same small difference between 1 W and 2 W as it does between 100 W and 200 W.

Changing decibels back to a power ratio (exponential form) is the same as for any logarithm with the addition of a multiplier, Fig. 7-2. The arrows in Fig. 7-2 indicate the transposition of quantities. Table 7-1 shows the number of decibels corresponding to various power ratios.

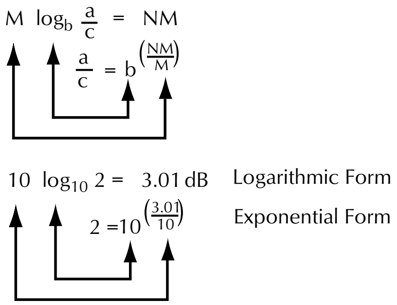

$M \log_b \frac{a}{c} = NM$
$\frac{a}{c} = b^{\left(\frac{NM}{M}\right)}$
$10 \log_{10} 2 = 3.01$ dB
Logarithmic Form
$2 = 10^{\left(\frac{3.01}{10}\right)}$
Exponential Form

Figure 7-2. Conversion of dB from logarithmic form to exponential form.

Finding Other Multipliers

Occasionally in acoustics, we may need multipliers other than 10 or 20. Once the ΔdB (the number of dB for a 2:1 ratio change) is known, calculate the multiplier by:

$$\text{log multiplier} = \frac{\log(Base) \times \Delta\text{dB}}{\log(Ratio)} \tag{7-7}$$

For example, if a 2:1 change is equivalent to 3.01 dB, then:

$$\begin{aligned}\text{log multiplier} &= \frac{\log(Base) \times 3.01}{\log 2}\\ &= 10\end{aligned}$$

or

$$10 \log 2 = 3.01$$

If a 2:1 change is equivalent to 6.02 dB, then:

$$\begin{aligned}\text{log multiplier} &= \frac{\log(Base) \times 6.02}{\log 2}\\ &= 20\end{aligned}$$

or

$$20 \log 2 = 6.02$$

Finally, if a 2:1 change is equivalent to 8 dB, then:

$$\begin{aligned}\textit{log multiplier} &= \frac{\log(Base) \times 8}{\log 2}\\ &= 26.58\end{aligned}$$

or

$$26.58 \log 2 = 8$$

For any ΔdB corresponding to a 2:1 ratio change involving logarithms to the base 10, this may be reduced to:

$$\textit{log multiplier} = 3.332 \times \Delta\text{dB} \tag{7-8}$$

Table 7-1. Power Ratios in Decibels

Power Ratio	Decibels (dB)
2	3.01030
3	4.77121
4	6.02060
5	6.98970
6	7.78151
7	8.45098

Table 7-1. (cont.) Power Ratios in Decibels

Power Ratio	Decibels (dB)
8	9.03090
9	9.54243
10	10.00000
100	20.00000
1000	30.00000
10,000	40.00000
100,000	50.00000
1,000,000	60.00000

7.3.4 The Decibel as a Power Quantity

We have seen that a number of decibels by themselves are only ratios. Given any reference (such as 50 W), we can use decibels to find absolute values. A standard reference for power in audio work is 10^{-3} W (0.001 W) or x V across $Z\Omega$. Note that when a level is expressed as a wattage, it is not necessary to state an impedance, but when it is stated as a voltage, an impedance is mandatory. This power is called 0 dBm. The small "m" stands for milliwatt (0.001 W) or one-thousandth of a watt.

Example

The power in watts corresponding to +30 dBm is calculated as follows:

$$10\log\frac{x}{0.001} = 30$$

$$\begin{aligned} \text{where } x &= 0.001 \times 10^{\frac{30}{10}} \\ &= 1\text{ W} \end{aligned}$$

For a power of −12 dBm:

$$0.001 \times 10^{\frac{-12}{10}} = 0.00006309\text{ W}$$

The voltage across 600 Ω is:

$$\begin{aligned} E &= \sqrt{WR} \\ &= \sqrt{0.00006309 \times 600} \\ &= 0.195\text{ V} \end{aligned}$$

Note that this −12 dBm power level can appear across any impedance and will *always* be the same power level. Voltages will vary to maintain this power level. In constant-voltage systems the power level varies as the impedance is changed. In constant-current systems the voltage changes as the impedance varies. (i.e.: −12 dBm across $8\,\Omega = \sqrt{0.00006309 \times 8} = 0.022\text{ V}$)

7.4 Measuring Electrical Power

$$W = EI\cos\theta \tag{7-9}$$

$$W = I^2 Z\cos\theta \tag{7-10}$$

$$W = \frac{E^2}{Z}\cos\theta \tag{7-11}$$

where,
W is the power in watts,
E is the electromotive force in rms volts,
I is the current in rms amperes,
Z is the magnitude of the impedance in ohms [in audio (ac) circuits Z (impedance) is used in place of R (ac resistance)],
θ is the phase difference between E and I in degrees.

These equations are only valid for single frequency rms sine wave voltages and currents.

Most Common Technique

1. Measure Z and θ.
2. Measure E across the actual load Z so that $W = (E^2/Z)\cos\theta$.

7.4.1 Expressing Power as an Audio Level

The reference power is 0.001 W (one milliwatt). When expressed as a level, this power is called 0 dBm (0 dB referenced to 1 mW).

Thus, to express a power level we need two powers—first the measured power W_1 and second the reference power W_2. This can be written as a power change in dB:

$$\begin{aligned} \frac{W_1}{W_2} &= \left(\frac{\dfrac{E_1^2}{1}}{\dfrac{E_2^2}{1}}\right)\left(\frac{\dfrac{1}{R_1}}{\dfrac{1}{R_2}}\right) \\ &= \left(\frac{E_1^2}{E_2^2}\right)\left(\frac{R_2}{R_1}\right) \end{aligned} \tag{7-12}$$

This can be written as a power level:

$$10\log\left[\left(\frac{E_1{}^2}{E_2{}^2}\right)\left(\frac{R_2}{R_1}\right)\right] = \text{power change in dB} \quad (7\text{-}13)$$

or

$$20\log\frac{E_1}{E_2} + 10\log\frac{R_2}{R_1} = \text{power change in dB} \quad (7\text{-}14)$$

Special Circumstance

When $R_1 = R_2$ and *only* then:

$$\text{Power level in dB} = 20\log\frac{E_1}{E_2} \quad (7\text{-}15)$$

where,

E_2 is the voltage associated with the reference power.

Conventional Practice

When calculating power level in dBm, we commonly make $E_2 = 0.775\,\text{V}$ and $R_2 = 600\,\Omega$. Note that E_2 may be any voltage and R_2 any resistance so long as together they represent 0.001 W.

7.5 Levels in dB

1. The term "level" is always used for a power expressed in decibels.
2. $10\log\frac{E_1{}^2}{E_2{}^2} = 10\log\frac{W_1}{W_2}$

   when $R_1 = R_2$

   $$2 \times 10\log\frac{E_1}{E_2} = 20\log\frac{E_1}{E_2}$$
   $$= 10\log\frac{W_1}{W_2}$$
3. Power definitions:

   Apparent power = $E \times I$ or E^2/Z,

   The average real or absorbed power is $(E^2/Z)\cos\theta$,

   The reactive power is $(E^2/Z)\sin\theta$,

   Power factor = $\cos\theta$.
4. The term "gain" or "loss" always means the power gain or power loss *at the system's output* due to the device under test.

7.5.1 Practical Variations of the dBm Equations

When the reference is the audio standard, i.e., 0.77459 V and 600 Ω, then:

$$\text{dB level to a reference} = 10\log\left[\left(\frac{E_1{}^2}{E_2{}^2}\right)\left(\frac{R_2}{R_1}\right)\right] \quad (7\text{-}16)$$

where,

$E_2 = 0.77459\ldots\text{V}$,

$R_2 = 600\,\Omega$.

then:

$$\frac{R_2}{E_2{}^2} = 1000$$

and 1/1000 = 0.001. Note that any E_2 and R_2 that result in a power of 0.001 W may be used. We can then write:

$$\text{Level (in dBm)} = 10\log\frac{E_1{}^2}{0.001R_1} \quad (7\text{-}17)$$

and

$$E_1 = \sqrt{0.001R_1\left(10^{\frac{dBm}{10}}\right)} \quad (7\text{-}18)$$

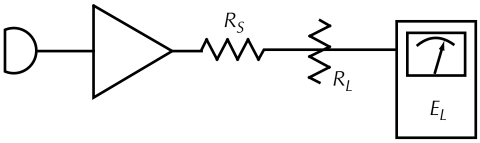
$$R_1 = \frac{E^2}{0.001\left(10^{\frac{dBm}{10}}\right)} \quad (7\text{-}19)$$

See Fig. 7-3.

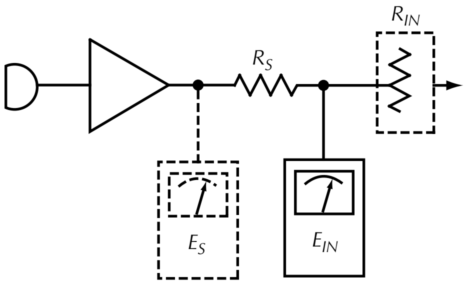
For all of the values in Table 7-2 the only thing known is the voltage. The indication is not a level. The apparent level can only be true across the actual reference impedance. Finally, the presence or absence of an attenuator or other sensitivity control is not known. See Section 7.20 for explanation of VU.

The power output of Boulder Dam is said to be approximately 3,160,000,000 W. Expressed in dBm, this output would be:

$$10\log\frac{3.16 \times 10^9}{10^{-3}} = 125\text{ dBm}.$$



A. Power across a load in dBm.



B. Available input power in dBm.

Figure 7-3. Power in dB across a load versus available input power.

Table 7-2. Root Mean Square Voltages Used as Nonstandard References

Voltage–V	Meter Indication	Apparent Level–VU	User
1.950	0	+8	Broadcast
1.230	0	+4	Recording
0.245	0	−10	Home recording
0.138	0	−15	Musical instruments

7.6 The Decibel in Acoustics—L_P, L_W, and L_I

In acoustics, the ratios most commonly encountered are changes in pressure levels. First, there must be a reference. The older level was 0.0002 dyn/cm^2, but this has recently been changed to 0.00002 N/m^2 (20 μN/m^2). Note that 0.0002 dyn/cm^2 is exactly the same sound pressure as 0.00002 N/m^2. Even more recently the standards group has named this same pressure pascals (Pa) and arranged this new unit so that:

$$20 \mu\text{Pa} = \frac{0.0002 \text{ dyn}}{\text{cm}^2} \tag{7-20}$$

This means that if the pressure is measured in pascals:

$$L_P = 20 \log \frac{x \text{ Pa}}{20 \ \mu\text{Pa}} \tag{7-21}$$

If the pressure is measured in dynes per square centimeter (dyn/cm^2), then:

$$L_P = 20 \log \frac{(x \text{ dyn/cm}^2)}{(0.0002 \text{ dyn/cm}^2)} \tag{7-22}$$

The root mean square sound pressure P can be found by:

$$P_{rms} = 2 \pi f A \rho c \tag{7-23}$$

where,
P_{rms} is in pascals,
f is the frequency in Hertz (Hz),
A is particle displacement in meters (rms value),
ρ is the density of air in kilograms per cubic meter (kg/m^3),
c is the velocity of sound in meters per second (m/s),
ρc = 406 RAYLS and is called the characteristic acoustic resistance (this value can vary),

or

$$L_P = 20 \log \frac{P_{rms}}{20 \ \mu\text{Pa}} \tag{7-24}$$

These are identical sound pressure levels bearing different labels. Sound pressure levels were identified as dB-SPL, and sound power levels were identified as dB-PWL. Currently, L_P is preferred for sound pressure level and L_W for sound power level. Sound intensity level is L_I:

$$L_I = 10 \log \frac{x \text{ W/m}^2}{10^{-12} \text{ W/m}^2} \tag{7-25}$$

At sea level, atmospheric pressure is equal to 2116 lb/ft^2. Remember the old physics lab stunt of partially filling an oil can with water, boiling the water, and then quickly sealing the can and putting it under the cold water faucet to condense the steam so that the atmospheric pressure would crush the can as the steam condensed, leaving a partial vacuum?

$$1 \text{ Atm} = 101{,}300 \text{ Pa}$$

therefore,

$$L_P = 20 \log \frac{101{,}300}{0.00002}$$
$$= 194 \text{ dB}$$

This represents the complete modulation of atmospheric pressure and would be the largest possible sinusoid. Note that the sound pressure (SP) is analogous to voltage. An L_P of 200 dB is the pressure generated by 50 lb of TNT at 10 ft. Table 7-3 shows the equivalents of sound pressure levels.

For additional insights into these basic relationships, the *Handbook of Noise Measurement* by Peterson and Gross is thorough, accurate, and readable.

Table 7-3. Equivalents of Pressure Levels

$$L_P = 20\log\frac{1\ \text{N/m}^2}{0.00002\ \text{N/m}^2}$$
$$= 93.98\ \text{dB}$$

Older values of a similar nature are:

1 microbar ≅ 1/1,000,000 of atmospheric pressure

≅ 74 dB

therefore,

1 Pa = 10 dyn/cm2

Other interesting figures:

Atmospheric pressure fully modulated $L_P \cong 194$ dB

1 lb/ft² $L_P = 127.6$ dB

1 lb/in² $L_P = 170.8$ dB

50 lb of TNT measured at 10 ft $L_P = 200$ dB

12 inch cannon, 12 ft in front of and below muzzle L_P = 220+ dB

Courtesy *GenRad Handbook*

7.7 Acoustic Intensity Level (L_I), Acoustic Power Level (L_W), and Acoustic Pressure Level (L_P)

7.7.1 Acoustic Intensity Level, L_I

The acoustic intensity I_a (the acoustic power per unit of area—usually in W/m² or W/cm² is found by:

$$L_I = 10\log\frac{x\text{W/m}^2}{10^{-12}\text{W/m}^2} \tag{7-26}$$

$$L_I = 10\log\frac{1.0\text{W/m}^2}{10^{-12}\ \text{W/m}^2}$$
$$= 120\ \text{dB}$$

7.7.2 Acoustic Power Level, L_W

The total acoustic power can also be expressed as a level (L_W):

$$L_W = 10\log\frac{\textit{Total acoustic watts}}{10^{-12}\ \text{W}} \tag{7-27}$$

7.7.3 Acoustic Pressure Level, L_P

To identify each of these parameters more clearly, consider a sphere with a radius of 0.282 m. (Since the surface area of a sphere equals $4\pi r^2$, this yields a sphere with a surface area of 1 m².) An omnidirectional point source radiating one acoustic watt is placed into the center of this sphere. Thus, we have, by definition, an acoustic intensity at the surface of the sphere of 1 W/m². From this we can calculate the P_{rms}:

$$P_{rms} = \sqrt{10W_a \times \rho c} \tag{7-28}$$

where,

W_a is the total acoustic power in watts,

ρc equals 406 RAYLS and is called the characteristic acoustic resistance.

Knowing the acoustic watts, P_{rms} is easy to find:

$$P_{rms} = \sqrt{10W_a \times 406}$$
$$= 20.15\ \text{Pa}$$

Thus, the L_P must be:

$$L_P = 20\log\frac{20.15\ \text{Pa}}{20\ \mu\text{Pa}}$$
$$= 120\ \text{dB}$$

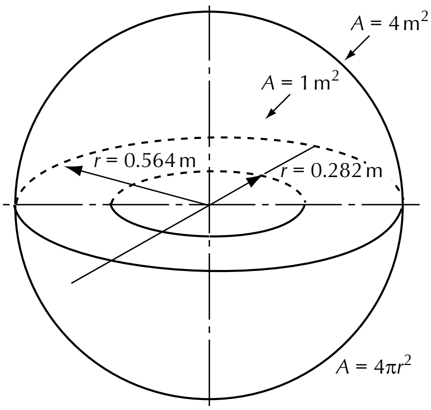
And the acoustic power level in L_W must be:

$$L_W = 10\log\frac{1\ \text{W}}{10^{-12}\ \text{W}}$$
$$= 120\ \text{dB}$$

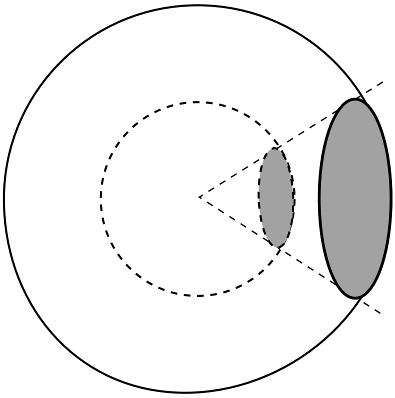
Thus, the L_P, L_I, and L_W at 0.282 m are the same numerical value if the source is omnidirectional, see Fig. 7-4.

7.8 Inverse Square Law

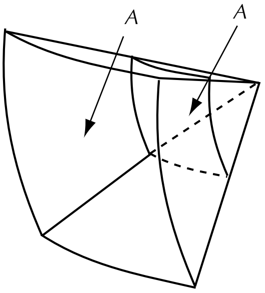
If we double the radius of the sphere to 0.564 m, the surface area of the sphere quadruples because the radius is squared in the area equation ($A = 4\pi r^2$). Thus, our intensity will drop to one-fourth its former value. (Note, however, that the total acoustic power is still 1 W so the L_W still is 120 dB.) Now an intensity change from 1 W to 0.25 W/m² can be written as a decibel change. The acoustic intensity (i.e., the power per unit of area), has dropped 6 dB in any given area:



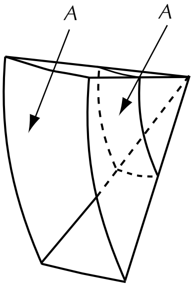
A. Sphere and radius.



B. Area increases with the square of the radius.



C. Area increases with the square of the radius when both angles diverge.



D Area increases as the radius increases when only one angle diverges.

Figure 7-4. Relationship of spherical surface area to radius.

$$L_I = 10\log\frac{0.25\ (\text{W/m}^2)(\text{new measurement})}{(1\ \text{W/m}^2)\left(\begin{matrix}\text{original reference}\\ \text{at the shorter radius}\end{matrix}\right)}$$

$$= -6.02\ \text{dB}$$

Therefore, our L_P had to also drop 6 dB and would now be approximately 114 dB.

This effect is commonly called the inverse square law change in level. Gravity, light, and many other physical effects exhibit this rate of change with varying distance from a source. Obviously, if you halve the radius, the levels all rise by 6 dB.

7.9 Directivity Factor

Finally, make the point source radiating one acoustic watt a hemispherical radiator instead of an omnidirectional one. Thus, at 0.282 m the surface area is now half of what our sphere had or 0.5 m². Therefore, our intensity is now 1 W/0.5 m² or the equivalent 2 W/m²:

$$10\log\frac{2\ \text{W}/\text{m}^2}{1\ \text{W}/\text{m}^2} = 3.01\ \text{dB}$$

Therefore, our L_P is 123.01 dB. L_w remains 120 dB. This 3.01 dB change represents a 2:1 change in the power per unit area; thus, a hemispherical radiator is said to have twice the directivity factor a spherical radiator has. Directivity factor is identified by a number of symbols—D_F, Q, $R\theta$, λ, M, etc. Q is the most widely used in the United States so we have chosen it for this text. (See Chapter 4 *Loudspeaker Directivity and Coverage* for a complete definition of Q.)

Directivity can also be expressed as a solid angle in steradians or $sr = 4\pi/Q$.

7.10 Ohm's Law

Recall that the use of the term "decibel" always implies a power ratio. Power itself is rarely measured as such. The most common quantity measured is voltage. If in measuring the voltage of a sinewave signal (oscillators are the most reliable and common of the test-signal sources) you obtain the root-mean square (rms) voltage, you can calculate the average power developed by using Ohm's law. Fig. 7-5 is a reminder of its many basic forms and uses the following definitions:

1. W is the average electrical power in watts (W).

2. I is the rms electrical current in amperes (A).
3. R is the electrical resistance in ohms (Ω).
4. E is the electromotive force in rms volts (V).
5. PF is the power factor ($\cos\theta$).

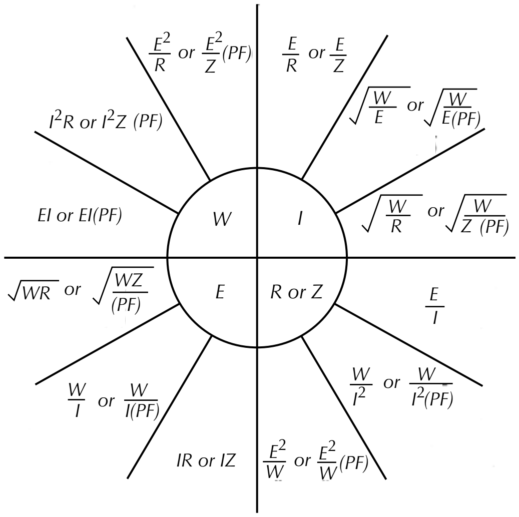

$\frac{E^2}{R}$ or $\frac{E^2}{Z}(PF)$
$\frac{E}{R}$ or $\frac{E}{Z}$
I^2R or I^2Z (PF)
$\sqrt{\frac{W}{E}}$ or $\sqrt{\frac{W}{E(PF)}}$
EI or $EI(PF)$
W
I
$\sqrt{\frac{W}{R}}$ or $\sqrt{\frac{W}{Z\ (PF)}}$
$\sqrt{WR}$ or $\sqrt{\frac{WZ}{(PF)}}$
E
R or Z
$\frac{E}{I}$
$\frac{W}{I}$ or $\frac{W}{I(PF)}$
$\frac{W}{I^2}$ or $\frac{W}{I^2(PF)}$
IR or IZ
$\frac{E^2}{W}$ or $\frac{E^2}{W}(PF)$

Figure 7-5. Ohm's law nomograph for ac or dc.

7.11 A Decibel Is a Decibel Is a Decibel

The decibel is always a power ratio; therefore, when dealing with quantities that are not power ratios, i.e., voltage, use the multiplier 20 in place of 10. As we encounter each reference for the dB, we will indicate the correct multiplier. Table 7-4 lists all the standard references, and Tables 7-5 through 7-8 contain additional information regarding reference labels and quantities. The decibel is not a unit of measurement like an inch, a watt, a liter, or a gram. It is the logarithm of a nondimensional ratio of two powerlike quantities.

Table 7-4. Common Decibel Notations and References

Quantity	Standard Reference	Symbol	Log Multiplier
Sound pressure	water: 1 dyn/cm^2 air: 0.0002 dyn/cm^2 or 0.00002 N/m^2	SPL or L_P	20
Sound intensity	10^{-16} W/cm^2 10^{-12} W/m^2		10
Sound power	10^{-12} W (new) 10^{-13} W (old)	PWL or L_w	10
Audio power	10^{-3} W	dBm	10
EMF	1 V	dBV	20
Amperes	1 mA		20
Acceleration	1 grms		20
Acceleration Spectral density	1 g^2/Hz		10
Volume units	10^{-3} W	VU	10
Distance	1 ft or 1 m	ΔD_x	20
Noise-ref	−90 dBm at 1 kHz	dBm	10

$$\text{dB} = \textit{Logarithm Multiplier} \times \log\frac{\textit{Quantity}}{\textit{Standard Reference}}$$

Table 7-5. Preferred Reference Labels for Acoustic Levels

Name	Definition
Sound pressure squared level	$L_P = 20\log(p/p_0)$ dB
Vibratory acceleration level	$L_a = 20\log(a/a_0)$ dB
Vibratory velocity level	$L_V = 20\log(v/v_0)$ dB
Vibratory force level	$L_F = 20\log(F/F_0)$ dB
Power level	$L_W = 10\log(P/P_0)$ dB
Intensity level	$L_I = 10\log(I/I_0)$ dB
Energy density level	$L_E = 10\log(E/E_0)$ dB

For $L_p = 20\log(x\ \text{Pa}/0.00002\ \text{Pa})$ use Eq.7-29.

$$L_P = (20\log x\ \text{Pa} + 94)\ \text{dB} \tag{7-29}$$

Table 7-6. A-Weighted Recommended Descriptor List

Term	Symbol
A-weighted sound level	L_A
A-weighted sound power level	L_{WA}
Maximum A-weighted sound level	L_{max}
Peak A-weighted sound level	L_{pk}
Level exceeded x% of the time	L_x
Equivalent sound level	L_{eq}
Equivalent sound level over time (T)	$L_{eq(T)}$
Day sound level	L_d
Night sound level	L_n
Day-night sound level	L_{dn}
Yearly day-night sound level	$L_{dn(Y)}$
Sound exposure level	L_{SE}

Table 7-7. Associated Standard Reference Values

1 atm = 1.013 bar = 1.033 kpa/cm2 = 14.70 lb/in^2 = 760 mm Hg = 29.92 in Hg

Acceleration of Gravity: g = 980.665 cm/s^2 = 32.174 ft/s^2 (standard or accepted value)

Sound Level: The common reference level is the audibility threshold at 1000 Hz, i.e., 0.0002 dyn/cm^2, 2×10^{-4} µbar, 2×10^{-5} N/m^2, 10^{-16} W/cm^2

Table 7-8. Recommended Descriptor List

Term	A-Weighting	Alternative* A-Weighting	Other Weighting†	Unweighted
Sound (pressure) level**	L_A	L_{pA}	L_B, L_{pB}	L_P
Sound power level	L_{WA}		L_{WB}	L_W
Max sound level	L_{max}	L_{Amax}	L_{Bmax}	L_{pmax}
Peak sound (pressure) level	L_{Apk}		L_{Bpk}	L_{pk}
Level exceeded x% of the time	L_x	L_{Ax}	L_{Bx}	L_{Px}
Equivalent sound level	L_{eq}	L_{Aeq}	L_{Beq}	L_{peq}
Equivalent sound level over time (T)‡	$L_{eq(T)}$	$L_{Aeq(T)}$	$L_{Beq(T)}$	$L_{peq(T)}$
Day sound level	L_d	L_{Ad}	L_{Bd}	L_{pd}
Night sound level	L_n	L_{An}	L_{Bn}	L_{pn}
Day-night sound level	L_{dn}	L_{Adn}	L_{Bdn}	L_{pdn}
Yearly day-night sound level	$L_{dn(Y)}$	$L_{Adn(Y)}$	$L_{Bdn(Y)}$	$L_{pdn(Y)}$
Sound exposure level	L_S	L_{SA}	L_{SB}	L_{Sp}
Energy average value over (nontime domain) set of observations	$L_{eq(e)}$	$L_{Aeq(e)}$	$L_{Beq(e)}$	$L_{peg(e)}$
Level exceeded x% of the total set of (nontime domain) observations	$L_{x(e)}$	$L_{Ax(e)}$	$L_{Bx(e)}$	$L_{px(e)}$
Average L_x value	L_x	L_{Ax}	L_{Bx}	L_{px}

*"Alternative" symbols may be used to assure clarity or consistency.

†Only B-weighting shown. Applies also to C, D, E weighting.

**The term "pressure" is used only for the unweighted level.

‡Unless otherwise specified, time is in hours (e.g., the hourly equivalent level is $L_{eq(1)}$). Time may be specified in nonquantitative terms (e.g., could be specified as $L_{eq(WASH)}$ to mean the washing cycle noise for a washing machine).

7.11.1 Older References

Much earlier, but valuable, literature used 10^{-13} W as a reference. In that case, the L_P value approximately equals the L_W value at 0.282 ft from an omnidirectional radiator in a free field (i.e., the number values are the same but, of course, different quantities are being measured). For 1 W using 10^{-12} W at 0.283 m, $L_W \cong L_P = 120$ dB. For 1 W using 10^{-13} W at 0.282 ft, $L_W \cong L_P = 130$ dB as found with the equation:

$$L_P = L_W - 10\log(4\pi r^2) \tag{7-30}$$

where,

L_W is 10log the wattage divided by the reference power 10^{-13},

r is the distance in meters from the center of the sound source.

Fig. 7-6 requires that you either know the distance from the source or assumes you are in the steady reverberant sound field of an enclosed space. L_P readings without one of these is meaningless. Fig. 7-7 shows typical power and L_W values for various acoustic sources.

7.12 The Equivalent Level (L_{EQ}) in Noise Measurements

Increasingly, acoustical workers in the noise control field are erecting an interesting edifice of measurement systems. A number of these measurement systems are based on the concept of average energy. Suppose, for example, that we have some means of collecting all of the A-weighted sound energy that arrives at a particular location over a certain period of time such as 90 dBA for 3.6 s (this could be a series of levels that lasted seconds, hours, or even days). We can then calculate the decibel level of steady noise for, say, 1 h that would be the equivalent level of the dBA for 3.6 s. That is, we wish to find the energy equivalent level for 1 h:

$$L_{EQ} = 10\log\left(\frac{1}{3600\,\text{s}}\int_0^{3.6\,\text{s}}\frac{P_A^{\,2}}{P_o^{\,2}}dt\right) \text{ in decibels} \tag{7-31}$$

where,

P_A is the acoustic pressure,

P_o is the reference acoustic pressure,

3600 s is the averaging time interval.

This integration reduces to

$$L_{EQ} = 10\log\left(\frac{10^{\frac{90}{10}} \times 3.6\text{ s}}{3600\text{ s}}\right)$$

Thus, 1.0 hour of noise energy at 60 dBA is the equivalent energy exposure of 90 dBA for 3.6 s.

L_{DN} (day-night level), *CNEL* (community noise level), etc., all follow similar schemes with variation in weightings for differing times of day, etc.

It is of interest that shooting a 0.458 magnum 174.7 L_P (peak) for 2.5 ms translates into:

$$L_{EQ} = 10\log\left(\frac{10^{\frac{174.7}{10}} \times 0.0025\text{ s}}{3600\text{ s}}\right)$$

$$= 113.12\text{ dB}$$

of steady sound for 1 h. OSHA allows only 15 min of exposure to levels of 110 dBA–115 dBA. As Howard Ruark's African guide, Harry Selby, remarked after Ruark had accidentally set off both barrels at once of a 0.470 express rifle while being charged by a Cape buffalo, "One of you ought to get up."

7.13 Combining Decibels

7.13.1 Adding Decibel Levels

The sum of two or more levels expressed in dB may be found as follows:

$$L_T = 10\log\left(10^{\frac{L_1}{10}} + 10^{\frac{L_2}{10}} + \ldots + 10^{\frac{L_N}{10}}\right) \qquad (7\text{-}32)$$

If, for example, we have a noisy piece of machinery with an $L_P = 90$ dB, and wish to turn on a second machine with an $L_P = 90$ dB, we need to know the combined L_P. Since both measured levels are the result of the power being applied to the machine, with some percentage being converted into acoustic power, we can determine LT by using Eq. 7-33. Therefore:

$$L_T = 10\log\left(10^{\frac{90}{10}} + 10^{\frac{90}{10}}\right)$$

$$= 10\log(10^9 + 10^9)$$

$$= 10\log(2 \times 10^9)$$

$$= 93\text{ dB}$$

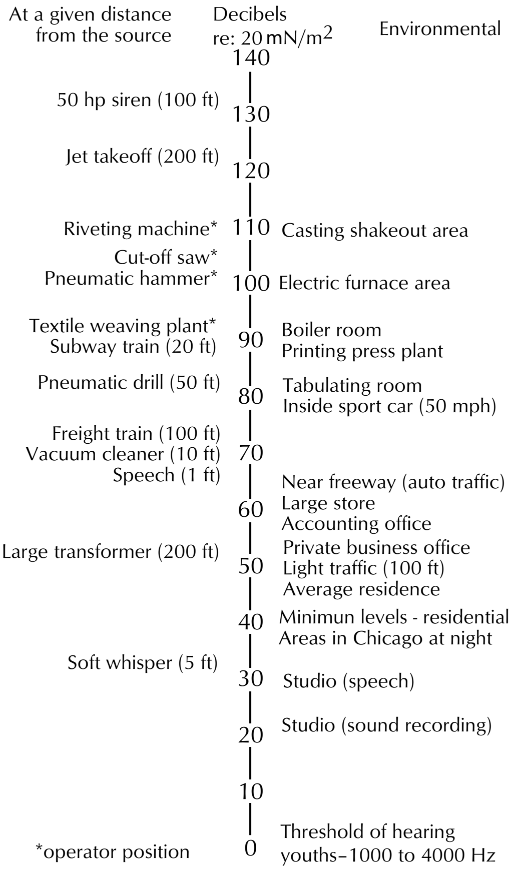

At a given distance from the source
Decibels re: 20 mN/m²
Environmental
140
50 hp siren (100 ft)
130
Jet takeoff (200 ft)
120
Riveting machine*
110
Casting shakeout area
Cut-off saw*
Pneumatic hammer*
100
Electric furnace area
Textile weaving plant*
Subway train (20 ft)
90
Boiler room
Printing press plant
Pneumatic drill (50 ft)
80
Tabulating room
Inside sport car (50 mph)
Freight train (100 ft)
Vacuum cleaner (10 ft)
70
Speech (1 ft)
Near freeway (auto traffic)
60
Large store
Accounting office
Large transformer (200 ft)
Private business office
50
Light traffic (100 ft)
Average residence
40
Minimun levels - residential
Areas in Chicago at night
Soft whisper (5 ft)
30
Studio (speech)
20
Studio (sound recording)
10
0
Threshold of hearing
youths–1000 to 4000 Hz
*operator position

Figure 7-6. Typical A-weighted sound levels as measured with a sound level meter. (*Courtesy GenRad*)

Doubling the acoustic power results in a 3 dB increase.

An alternative dB addition technique is given through the courtesy of Gary Berner.

$$L_T = 10\log\left(10^{\frac{-(diff\ in\ dB)}{10}} + 1\right) + smallest\ number \qquad (7\text{-}33)$$

Example

If we wish to add 90 dB to 96 dB, using Eq. 7-33, take the difference in dB (6 dB) and put it in the equation:

$$L_T = 10\log\left(10^{\frac{6}{10}} + 1\right) + 90$$

$$= 96.97\text{ dB}$$

Input signals to a mixing network also combine in this same manner, but the insertion loss of the

Power (watts)	Power Level dB re: 10^{-12}W	Source	
20 to 40 million	195	Saturn rocket	
100,000	170	Ram jet	
		Turbojet engine with afterburner	
10,000	160	Turbojet engine (7000 lb thrust)	
1000	150		
		4 propeller airliner	
100	140		
10	130	75-piece orchestra Pipe organ	Peak rms values in $1/8$ s intervals
		Small aircraft engine	
1	120	Large chipping hammer	
		Piano BB*b* tuba	Peak rms values in $1/8$ s intervals
0.1	110	Blaring radio	
		Centrifugal ventilating fan (13,000 CFM)	
0.01	100	4-ft loom	
		Auto on highway	
0.001	90	Vane axial ventilating fan (1500 CFM) Voice—shouting (average long term rms)	
0.0001	80		
0.000 01	70	Voice—conversational level (average long-time rms)	
0.000 001	60		
0.000 000 1	50		
0.000 000 01	40		
0.000 000 001	30	Voice—very soft wisper	

Figure 7-7. Typical power and L_W values for various acoustic sources.

network must be subtracted. Two phase-coherent sinewave signals of equal amplitude will combine to give a level 6 dB higher than either sinewave.

The general case equation for adding either sound pressure, voltages, or currents is:

$$\text{Combined } L_P = 20\log\sqrt{\left(10^{\frac{E_1}{20}}\right)^2+\left(10^{\frac{E_2}{20}}\right)^2+2\left(10^{\frac{E_1}{20}}\right)\left(10^{\frac{E_2}{20}}\right)(\cos[a_1-a_2])} \qquad (7\text{-}34)$$

Table 7-9 shows the effects of adding two equal amplitude signals with different phase together using Eq. 7-36.

7.13.2 Subtracting Decibels

The difference of two levels expressed in dB may be found as follows:

$$L_{diff} = 10\log\left(10^{\frac{Total\ Level}{10}} - 10^{\frac{Level\ with\ one\ source\ off}{10}}\right) \qquad (7\text{-}35)$$

7.13.3 Combining Levels of Uncorrelated Noise Signals

When the sound level of a source is measured in the presence of noise, it is necessary to subtract out the effect of the noise on the reading. First, take a reading of the source and the noise combined (L_{S+N}). Then take another reading of the noise alone (the source having been shut off). The second reading is designated L_N. Then:

$$L_S = 10\log\left(10^{\frac{L_{S+N}}{10}} - 10^{\frac{L_N}{10}}\right) \qquad (7\text{-}36)$$

To combine the levels of uncorrelated noise signals we can also use the chart in Fig. 7-8 as follows

Table 7-9. Combining Pure Tones of the Same Frequency but Differing Phase Angles

Signal 1 Amplitude, L_P in dB	Signal 1 Phase, in Degrees	Signal 2 Amplitude, L_P in dB	Signal 2 Phase, in Degrees	Combined Signal Amplitude, L_P in dB
90	0	+90	0	96.02
90	0	+90	10	95.99
90	0	+90	20	95.89
90	0	+90	30	95.72
90	0	+90	40	95.48
90	0	+90	50	95.17
90	0	+90	60	94.77
90	0	+90	70	94.29
90	0	+90	80	93.71
90	0	+90	90	93.01
90	0	+90	100	92.18
90	0	+90	110	91.19
90	0	+90	120	90.00
90	0	+90	130	88.54
90	0	+90	140	86.70
90	0	+90	150	84.28
90	0	+90	160	80.81
90	0	+90	170	74.83
90	0	+90	180	$-\infty$

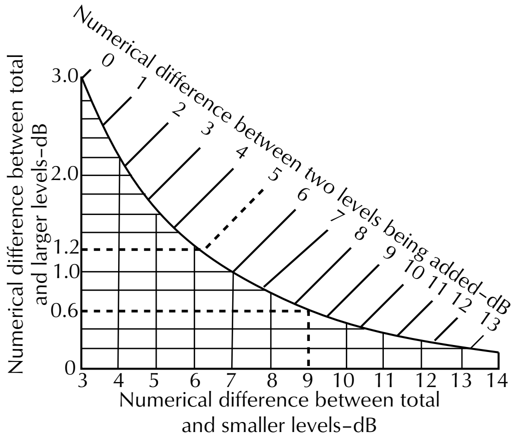


Figure 7-8. Chart used for determining the combined level of uncorrelated noise signals.

7.13.4 To Add Levels

Enter the chart with the numerical difference between the two levels being added (top of chart). Follow the line corresponding to this value to its intersection with the curved line; then move left to read the numerical difference between the total and larger levels. Add this value to the larger level to determine the total.

Example

To add 75 dB to 80 dB, subtract 75 dB from 80 dB; the difference is 5 dB. In Fig. 7-8, the 5 dB line intersects the curved line at 1.2 dB on the vertical scale. Thus the total value is 80 dB + 1.2 dB, or 81.2 dB.

7.13.5 To Subtract Levels

Enter the chart in Fig. 7-8 with the numerical difference between the total and larger levels if this value is less than 3 dB. Enter the chart with the numerical difference between the total and smaller levels if this value is between 3 dB and 14 dB. Follow the line corresponding to this value to its intersection with the curved line, then either left or down to read the numerical difference between total and larger (smaller) levels. Subtract this value from the total level to determine the unknown level.

Example

Subtract 81 dB from 90 dB; the difference is 9 dB. The 9 dB vertical line intersects the curved line at 0.6 dB on the vertical scale. Thus the unknown level is 90 dB – 0.6 dB, or 89.4 dB.

7.14 Combining Voltage

To combine voltages, use the following equation:

$$E_T = \sqrt{E_1^2 + E_2^2 + 2E_1E_2[\cos(a_1 - a_2)]} \tag{7-37}$$

where,

E_T is the total sound pressure, current, or voltage,

E_1 is the sound pressure, current, or voltage of the first signal,

E_2 is the sound pressure, current, or voltage of the second signal,

a_1 is the phase angle of signal one,

a_2 is the phase angle of signal two.

7.15 Using the Log Charts

7.15.1 The 10Log *x* Chart

There are two scales on the top of the $10\log_{10} x$ chart in Fig. 7-9. One is in dB above and below a 1 W reference level, and the other is in dBm (reference 0.001 W). Power ratios may be read directly from the 1 W dB scale.

Example

How many decibels is a 25:1 power ratio?

1. Look up 25 on the Power–watts scale.
2. Read 14 dB directly above the 25.

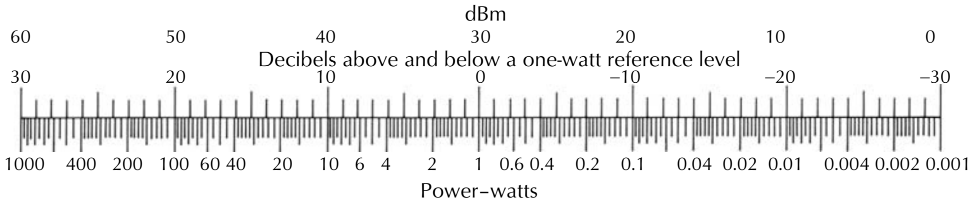


Figure 7-9. The $10\log_{10} x$ chart.

Example

We have a 100 W amplifier but plan to use a 12 dB margin for "head room." How many watts will our program level be?

1. Above 100 W find +50 dBm.
2. Subtract 12 dB from 50 dBm to obtain +38 dBm. Just below +38 dBm find approximately 6 W.

Example

A 100 W amplifier has 64 dB of gain. What input level in dBm will drive it to full power?

1. Above 100 W read +50 dBm.
2. +50 dBm – 64 dB gain = –14 dBm.

Example

A loudspeaker has a sensitivity of L_P = 99 dB at 4 ft with a 1 W input. How many watts are needed to have an L_P of 115 at 4 ft?

1. 115 L_P – 99 L_P = +16 dB.
2. At +16 on the 1 W scale read 39.8 W.

7.15.2 The 20Log x Chart

Refer to the chart in Fig. 7-10. A 2:1 voltage, distance, or sound pressure change is found by locating 2 on the ratio or D scale and looking directly above to 6 dB.

Example

A loudspeaker has a sensitivity of L_P = 99 dB at 4 ft with 1 W of input power. What will the level be at 100 ft?

1. Find the relative dB for 4 ft (dB_{Rel} = 12 dB).
2. Find the relative dB for 100 ft (dB_{Rel} = 40 dB).
3. Calculate the absolute dB (40 dB – 12 dB = 28 dB).
4. L_P = 99 dB – 28 dB = 71 dB.

Example

If we raise the voltage from 2 V to 10 V, how many decibels would we increase the power?

1. Find the relative dB for a ratio of 2 (dB_{Rel} = 6 dB).
2. Find the relative dB for a ratio of 10 (dB_{Rel} = 20 dB).
3. Absolute dB change = 20 dB – 6 dB = 14 dB.
4. Since a dB is a dB, the power also changed by 14 dB.

7.16 Finding the Logarithm of a Number to Any Base

In communication theory, the base 2 is used. Occasionally other bases are chosen. To find the logarithm of a number to any possible given base, write:

$$x = b^n \tag{7-38}$$

where,
x is the number for which a logarithm is to be found,
b is the base,
n is the logarithm.

Then write:

$$\log x = n \log b \tag{7-39}$$

and

$$\frac{\log x}{\log b} = n \tag{7-40}$$

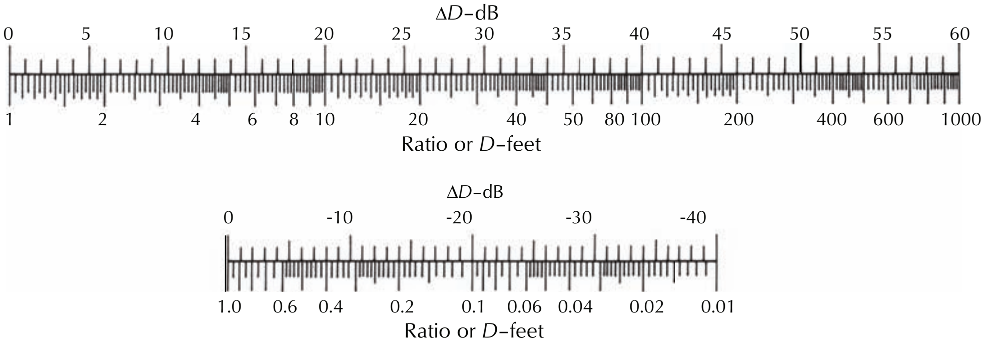


Figure 7-10. The $20\log_{10} x$ chart.

Suppose we want to find the natural logarithm of 2 (written ln2). The base of natural logarithms is $e = 2.7188281828$. Then:

$$\frac{\log 2}{\log e} = \frac{0.30103}{0.43425} = 0.69315$$

To verify this result:

$$e^{0.69315} = 2$$

To find $\log_2$ of 26:

$$\frac{\log 26}{\log 2} = \frac{1.41497}{0.30103} = 4.70044$$

The general case is:

$$\frac{\log_{10} \textit{of the number}}{\log_{10} \textit{of the base}} = \log_{\text{base}} \textit{of the number} \quad (7\text{-}41)$$

7.17 Semitone Intervals

Suppose that we need $\sqrt[12]{2}$ (the semitone interval in music). We could write:

$$\frac{\log 2}{12} = \log \sqrt[12]{2} \quad (7\text{-}42)$$

Therefore,

$$10^{\frac{\log 2}{12}} = 10^{\frac{0.30}{12}} = 10^{0.02508} = 1.05946 = \sqrt[12]{2}$$

This is the same as multiplying 1.05946 by itself 12 times to obtain 2.

$10^{0.02508}$ is called the antilog of 0.02508. The antilog is also written as $\log^{-1}$, antilog 10, or 10 exp. All these terms mean exactly the same thing.

7.18 System Gain Changes

Imagine a noise generator driving a power amplifier and a loudspeaker, Fig. 72-11. If the voltage out of the noise generator is raised by 6 dB, what happens?

Voltage	Electrical Power	L_P	L_W
Doubled	Quadrupled	Doubled	Quadrupled
+6 dB	+6 dB	+6 dB	+6 dB

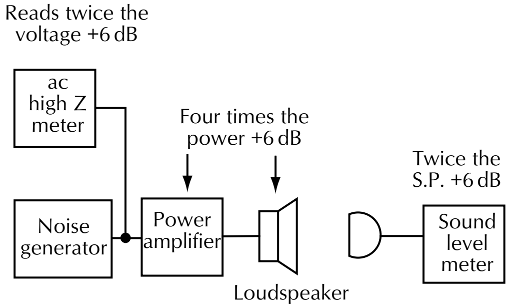
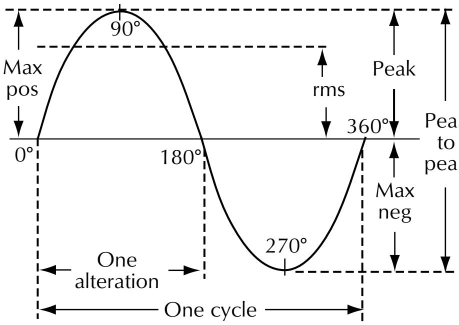


Figure 7-11. Voltage, electrical power, P_w, and sound pressure, compared.

This means that, in a linear system, a level change ahead of any components results in a level change for that same signal in all subsequent components, though it might be measured as quite different voltages or wattages at differing points. The change in level at any point would be the same. We will work with this concept a little later when we plot the gains and losses through a total system.

7.19 The VU and the VI Instrument

Volts, amperes, and watts can be measured by inserting an appropriate meter into the circuit. If all audio signals were sinewaves, we could insert a dBm meter into the circuit and get a reading that would correlate with both electrical and acoustical variations. Unfortunately, audio signals are complex waveforms, and their rms value is not 0.707 times peak but can range from as small as 0.04 times peak to as high as 0.99 times peak, Fig. 7-12. To solve this problem, broadcasting and telephone engineers got together in 1939 and designed a special instrument for measuring speech and music in communication circuits. They calibrated this new type of instrument in units called VU. The dBm and the VU are almost identical, the only difference being in their usage. The instrument used to measure VU is called the volume indicator (VI) instrument. (Some users ignore this and incorrectly call it a VU meter.) Both dBm meters and volume indicator instruments are specially calibrated voltmeters. Consequently, the VU and dBm scales on these meters give correct readings only when the measurement is being made across the impedance for which they are calibrated (usually 150 Ω or 600 Ω). Readings taken across the design impedance are referred to as true levels, whereas readings taken across other impedances are called apparent levels.



rms = 0.707 × peak voltage
rms = 0.3535 × peak to peak voltage
peak = 1.414 × rms voltage
peak-to-peak = 2.828 × rms voltage

Figure 7-12. Sinewave voltage values. The average voltage of a sine wave is zero.

Apparent levels can be useful for relative frequency response measurements, for example. When the impedance is not 600 Ω, the correction factor of 10log(600/*new impedance*) can be added to the formula containing the reference level as in the following equation:

$$True\ VU = Apparent\ VU + 10\log\frac{600}{Z\ measured} \qquad (7\text{-}43)$$

The result is the true level.

7.19.1 Crest Factor

The crest factor (CF) is the ratio of the peak output to the average output. It is typically graphed in terms of the output power and is expressed in dB. For example, the CF of a sine wave is 3 dB. The CF of music may vary between 6 dB and 24 dB.

Crest factor is defined as ten times the logarithm of peak power out divided by average power. The actual measurements are often made using voltage. In that case we divide the peak voltage squared by the rms voltage squared. The root mean square integral is;

$$e_{rms} = \left(\frac{1}{T}\int_{0e}^{T} e^2 dt\right)^{0.5} \qquad (7\text{-}44)$$

The rms voltage squared divided by the impedance is average power.

Using voltage implies that both measurements were made across identical impedances or (open circuit). Crest factor applied to voltage waveforms is a common practice but not the defined term. Powers are chosen because when impedances vary widely over the bandwidth of interest, and consequently the bandwidth is divided into components, the powers can be added to obtain the overall crest factor. The Texas Instruments "*Guidelines for Measuring Audio Power Amplifier Performance*" page 25 displays actual measurements and the usefulness of power measurements. When we remember that power describes heat dissipation in power amplifiers then CF as a ratio of powers becomes evident.

7.19.2 The VU Impedance Correction

When a VI instrument is connected across 600 Ω and is indicating 0 VU on a sinewave signal, the true level is 4 dB higher, or +4 dBm, instead of 0 dBm or zero level. The reason this is so is shown in Fig. 7-13. The VI instrument uses a 50 μA

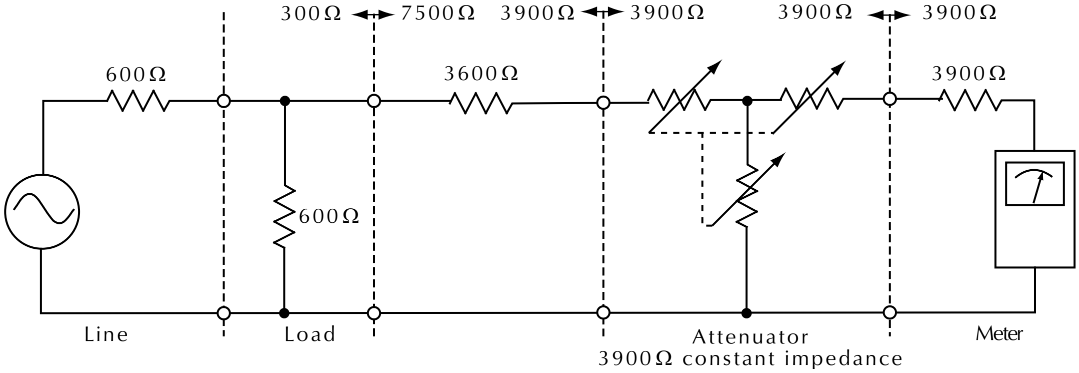
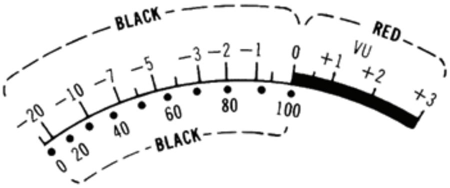


Figure 7-13. VI instrument circuit.

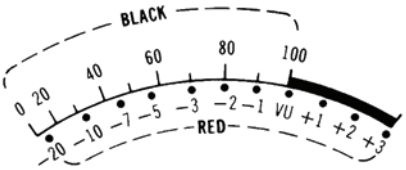
D'Arsonval movement in conjunction with a copper-oxide bridge-type rectifier. The impedance of the instrument and rectifier is 3900 Ω. To minimize its effect when placed across a 600 Ω line, it is "built-out" an additional 3600 Ω to a total value of 7500 Ω. The addition of this build-out resistance causes a 4 dB loss between the circuit being measured and the instrument. Therefore, when a properly installed VI instrument is fed with 0 dBm across a 600 Ω line, the meter would actually read –4 VU on its scale. (When the attenuator setting is added, the total reading is indeed 0 VU.)

Presently, no major U.S. manufacturer offers for sale a standard volume indicator that complies with the applicable standard (C16.5). The standard requires that an attenuator be supplied with the instrument and none of the manufacturers do so. What they are doing requires some attention. The instruments (usually high-impedance bridge-types) are calibrated so as to act as if the attenuator were present. When the meter reads 0 VU (on a sinewave for calibration purposes), the true level is +4 dBm. This means a voltage of 1.23 V across 600 Ω will cause the instrument to read an apparent 0 VU. Note that when reading sine wave levels, the label used is "dBm." When measuring program levels, the label used is "VU." The VU value is always the instrument indication plus the attenuator value.

Two different types of scales are available for VI meters, Fig. 7-14. Scale A is a VU scale (recording studio use), and scale B is a modulation scale (broadcast use). On complex waveforms (speech and music), the readings observed and the peak levels present are about 10 dB apart. This means that with a mixer amplifier having a sinewave output capability of +18 dBm, you are in danger of distortion with any signal indicating more than +8 VU on the VI instrument (+18 dBm – [+10 dB] peaking factor or meter lag equals +8 VU).



A. Recording and test equipment.



B. Broadcast monitoring.

Figure 7-14. VI instrument scales.

Fig. 7-15 shows an example of commercially available VI instrument panels used in the past that included the VI instrument and 3900 Ω attenuator, which also contains the 3600 Ω build-out resistor.

Figure 7-15. Examples of commercial-type VI instrument panels.

7.19.3 How to Read the VU Level on a VI Instrument

A VI instrument is used to measure the level of a signal in VU. In calibration: 0 VU = 0 dBm and a 1.0 VU increment is identical to a 1.0 dB increment. The true level reading in VU is found by:

$$\text{True VU level} = \text{Apparent level} + \text{Impedance correction} \tag{7-45}$$

or

$$\text{True VU level} = \text{Instrument indication} + 10\log\left(\frac{600}{Z_{act}}\right)$$

where,

Apparent level = Instrument indication + Attenuator or sensitivity indicator.

Thus, we can have:

1. A direct reading from the face of the instrument (zero preferred).
2. The reading from the face of the instrument plus the reading from the attenuator or other sensitivity adjustment—normally a minimum of +4 dB or higher. When the instrument indicates zero, the apparent level is the attenuator setting.
3. The correction factor for impedance other than the reference impedance. 600 Ω is the normal impedance chosen for a reference, but any value

can be used so long as the voltage across it results in 0.001 W, Fig. 7-16.

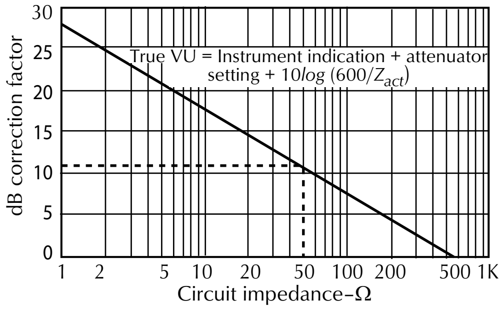


Figure 7-16. Relationship between circuit impedance and the dB correction value.

Example

We have an indication on the instrument of –4 VU. The sensitivity control is at +4. We are across 50 Ω (a 100 W amplifier 70.7 V output). Using Fig. 7-16, our true VU would be –4 VU + (+4 VU) + 10.8 correction factor = 10.8 VU.

7.19.4 Calibrating a VI Instrument

The instrument should be calibrated to read a true level of zero VU when an input of a 1000 Hz steady-state sinewave signal of 0 dBm (0.001 W) is connected to it. For example, typical calibration is when the instrument indicates –4, the attenuator value is +4, and it is connected across a 600 Ω circuit. Levels read on a VI instrument when the source is the aforementioned sinewave signal should be stated as dBm levels.

7.19.5 Reading a VI Instrument on Program Material

Because of the ballistic properties of VI instruments, they exhibit what has been called "instrument lag." On short-duration peak levels, they will "lag" by approximately 10 dB. Stated another way, if we read a true VU level of +8 VU on a speech signal, then the level in dBm becomes +18 dBm. This means that the associated amplification equipment, when fed a true VU level of +8 VU, must have a steady-state sinewave capability of +18 dBm to avoid overload.

Rule

Levels stated in VU are assumed to be program material and levels stated in dBm are assumed to be steady-state sinewave.

7.19.6 Reading Apparent VU Levels

VI instruments can be used to read apparent or relative levels. If, for example, you know that overload occurs at some apparent level. You can use that reading as a satisfactory guide to the system's operation, even though you do not know the true level. When adjusting levels using the instrument to read the relative change in level, such as turning the system down 6 dB, you do not need to do so in true level readings. Instrument indication serves effectively in such cases.

When being given a level, be sure to ascertain whether it is:

1. An instrument indication.
2. An apparent level.
3. A true level.
4. A relative level.
5. A calibration level.
6. A program level.
7. None of the above but simply an arbitrary meter reading.

Special Note: Well-designed mixers have instruments that indicate the available input power level to the device connected to its output. Such levels are true levels.

7.20 Calculating the Number of Decades in a Frequency Span

To find the relationship of the number of decades between the lowest and highest frequencies, use the following equations:

$$\frac{H.F.}{L.F.} = 10^1 = 1\text{ decade} \qquad (7\text{-}46)$$

therefore:

$$\begin{aligned} \frac{H.F.}{L.F.} &= 10^{x\text{ decade}} \\ \frac{\ln H.F. - \ln L.F.}{\ln 10} &= x\text{ decades} \end{aligned} \qquad (7\text{-}47)$$

or

$$\ln H.F. - \ln L.F. = \ln 10 \times (x \text{ decades}) \tag{7-48}$$

Further:

$$H.F. = e^{(x \text{ decades}) \times (\ln 10) + \ln L.F.} \tag{7-49}$$

and

$$L.F. = e^{[\ln H.F. - (x \text{ decades} \times \ln 10)]} \tag{7-50}$$

Example

How many decades does the bandpass 500 Hz to 12,500 Hz contain? Using Eq. 7-48:

$$\frac{\ln 12{,}500 - \ln 500}{\ln 10} = 1.39794 \text{ decades}$$

If we had 12,500 Hz as a *H.F.* limit and wished to know the low frequency that would give us 1.4 decades, we would calculate:

$$\begin{aligned} L.F. &= e^{[\ln 12{,}500 - (1.4 \text{ decades} \times \ln 10)]} \\ &= 497.63 \text{ Hz} \end{aligned}$$

If we had the *L.F.* limit and wished to know the *H.F.*, then:

$$\begin{aligned} H.F. &= e^{(1.4 \text{ decades} \times \ln 10) + \ln 497.63} \\ &= (12{,}500 \text{ Hz}) \end{aligned}$$

7.21 Deflection of the Eardrum at Various Sound Levels

If we make the assumption that the eardrum displacement is the same as that of the air striking it we can write:

$$D_{in} = 3 \times 10^{-7}\left(\frac{10^{\frac{L_P}{20}}}{f}\right) \tag{7-51}$$

or

$$D_{cm} = 39 \times 10^{-3}\left(\frac{0.0002 \times 10^{\frac{L_P}{20}}}{f}\right) \tag{7-52}$$

where,

D_{in} is the displacement in inches (the rms amplitude) of the air,

D_{cm} is the displacement in cm,

f is the frequency in Hz,

L_P is the sound level in decibels referred to 0.00002 N/m^2.

Example

What is the displacement of the eardrum in inches for a tone at 1000 Hz at a level of 74 dB? Using Eq. 7-51:

$$\begin{aligned} D_{in} &= 3 \times 10^{-7}\left(\frac{10^{\frac{74}{20}}}{1000}\right) \\ &= 0.0000015 \text{ in} \end{aligned}$$

which is a displacement of approximately one-one-millionth of an inch (0.000001 in).

7.22 The Phon

Fig. 7-17 shows free-field equal-loudness contours for pure tones (observer facing source), determined by Robinson and Dadson at the National Physical Laboratory, Teddington, England, in 1956 (ISO/R226-1961). The phon scale is of equal loudness level contours. At 1000 Hz every decibel is the equivalent loudness of a phon unit.

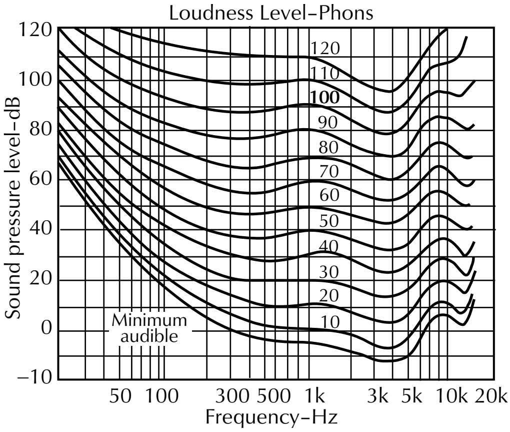


Figure 7-17. Equal loudness contours.

For two different sounds within a critical band (for most practical purposes, using ⅓ octave bands suffices) they are added in the same manner as decibel readings.

$$P_T = 10\log\left(10^{\frac{L_{P1}}{10}} + 10^{\frac{L_{P2}}{10}}\right) \tag{7-53}$$

$$= \text{phons}$$

where,

L_{P1} and L_{P2} are the individual sound levels in dB.

For example, suppose that within the same critical band we have two tones each at 70 phons. Using Eq. 7-53:

$$P_T = 10\log\left(10^{\frac{70}{10}} + 10^{\frac{70}{10}}\right)$$

$$= 73 \text{ phons}$$

An interesting experiment in this regard is to start with two equal level signals 10 Hz apart at 1000 Hz and gradually separate them in frequency while maintaining their phon level.

They will increase in apparent loudness as they separate. This is one of the reasons a distorted system sounds louder than an undistorted system at equal power levels. One final factor worthy of storage in your own mental "read only memory" is that in the 1000 Hz region most listeners judge a change in level of 10 dB as twice or half the loudness of the original tone.

Fig. 7-18 is a chart of frequency and dynamic range for various musical instruments and the upper and lower frequency range of the average young adult.

7.23 The Tempered Scale

The equal tempered musical scale is composed of 12 equally spaced intervals separated by a factor of $\sqrt[12]{2}$. All notes on the musical scale (excluding sharps and flats) however, are not equally spaced. This is because there are two ½ step intervals on the scale: that between E and F, and that between B and C. The 12 tones, therefore, go as follows: C, C#, D, D#, E, F, F#, G, G#, A, A#, B, C, see Table 7-10.

Table 7-10. Tempered Scale

Note	Frequency Ratio	Frequency–Hz
C	1.000	262
C#, D♭	1.059	277
D	1.122	294
D#, E♭	1.189	311
E	1.260	330
F	1.335	349
F#, G♭	1.414	370
G	1.498	392
G#, A♭	1.587	415
A	1.682	440
A#, B♭	1.782	466
B	1.888	494
C	2.000	523

7.24 Measuring Distortion

Fig. 7-19 illustrates one of the ways of measuring harmonic distortion. Two main methods are employed. One uses a band rejection filter of narrow bandwidth having a rejection capability of at least 80 dB in the center of the notch. This deep notch "rejects" the fundamental of the test signal (usually a known-quality sinewave from a test audio oscillator) and permits reading the noise voltage of everything remaining in the rest of the bandpass. Unfortunately, this also includes the hum and noise as well as the harmonic content of the equipment being tested, Fig. 7-20.

The second method is more useful. It uses a tunable wave analyzer. This instrument allows the measurement of the amplitudes of the fundamental and each harmonic, as well as identifying the hum, amplitude, and the noise spectrum shape, Fig. 7-20. Such analyzers come in many different bandwidths, with a ⅒ octave unit allowing readings down to 1% of the fundamental (it is –45 dB at 2*f*). By looking at Fig. 7-20, it is easy to see that harmonic distortion appears as a spurious noise. Today tracking filter wave analysis allows nonlinear distortion behavior to be "tracked" or measured.

7.25 The Acoustical Meaning of Harmonic Distortion

The availability of extremely wide-band amplifiers with distortions approaching the infinitesimal and the gradual engineering of a limited number of loudspeakers with distortions just under 1% at usable levels (90–100 dB *SPL* at 10–12 ft) brings up an interesting question: "How low a distortion is really needed?"

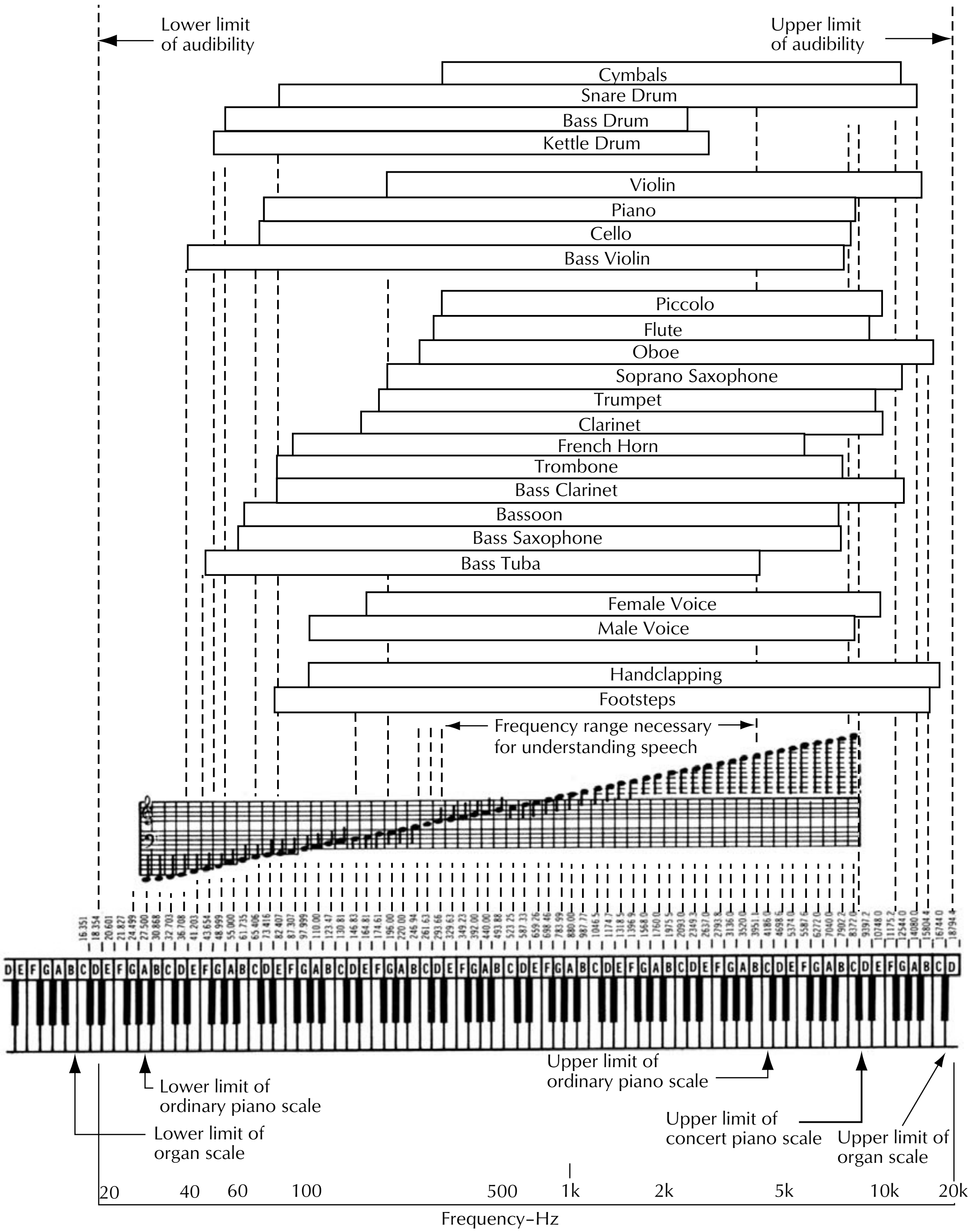


Figure 7-18. Audible frequency range.

7.25.1 Calculating the Maximum Allowable Total Harmonic Distortion in an Arena Sound System

The most difficult parameter to achieve in the typical arena sound system is sufficient signal-to-noise ratio (*SNR*) to ensure acceptable articulation losses for consonants in speech. It must be at least 25 dB. In that case, the total harmonic distortion should be at least 10 dB below the 25 dB

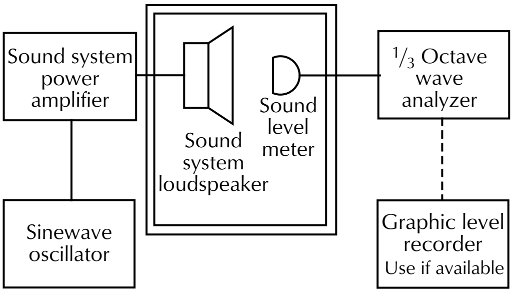


Figure 7-19. Measurement of harmonic distortion.

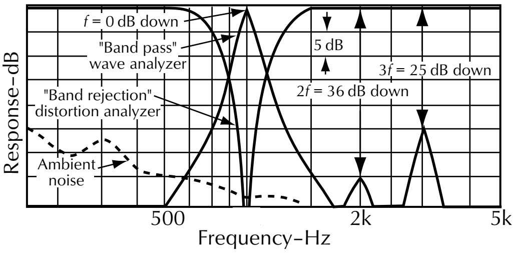


Figure 7-20. Methods of measuring distortion.

SNR to avoid the addition of the two signals. If both signals were at the same level, a 3 dB increase in level would occur. Therefore, (–25 dB) + (–10 dB) means that the total harmonic distortion (*THD*) should not exceed –35 dB.

$$Percentage = 100 \times 10^{\frac{\pm dB}{20}} \tag{7-54}$$

Therefore, we could calculate:

$$100 \times 10^{\frac{-35}{20}} = 1.78\%$$

This is why carefully thought-out designs for use in heavy-duty commercial sound work have a *THD* of 0.8 to 0.9%:

$$20\log\frac{100 \pm x\%}{100} = \text{dB change}$$

Since the 0.8% already represents (100 – 99.2), we can write

$$20\log\frac{0.8}{100} = -42\ \text{dB}$$

Now, suppose an amplifier has 0.001% distortion. What sort of dynamic range does this represent?

$$20\log\frac{0.001}{100} = -100$$

That is a power ratio of:

$$10^{\frac{100}{10}} = 10{,}000{,}000{,}000$$

We can conclude that if such a figure were achievable, it would nevertheless not be useful in arena systems.

7.26 Playback Systems in Studios

Assume that a monitor loudspeaker can develop L_P = 110 dB at the mixer's ears and that in an exceptionally quiet studio we reach L_P = 18 dB at 2000 Hz (NC-20). We then have

$$L_{P_{Diff}} = L_{P_{Total}} - L_{P_{Noise}} \tag{7-55}$$

which is equal to 92 dB. Adding 10 dB to avoid the inadvertent addition of levels gives 102 dB. The distortion now becomes:

$$100 \times 10^{\frac{-102}{20}} = 0.00078\%$$

In this case, extraordinary as it is, the previously esoteric figure becomes a useful parameter.

7.26.1 Choosing an Amplifier

As we pointed out earlier, the loudspeaker will establish equilibrium around 1% with its acoustic distortion. To the builder of systems, this means that extremely low distortion figures cannot be used within the system as a whole. Therefore, systems-oriented amplifier designers have not attempted to extend the bandpass to extreme limits. They know that they must balance bandpass, distortion, noise, and hum against stability with all types of loads, extensions of mean time-before-failure characteristics. Most high quality sound reinforcement amplifiers incorporate an output transformer, giving us 70 V, 25 V, and 4 Ω, 8 Ω, and 16 Ω outputs. In fact, connecting across the 4 Ω and 8 Ω taps yields a 0.69 Ω output.

Example:

Let the rms speech value be L_P = 65 dB at 2 ft in the 1000–2000 Hz octave band, Fig. 7-21. Let the ambient noise level be L_P = 32 dB with the air conditioning on and 16 dB with the air conditioning off in the same octave band, Fig. 7-22. With the air conditioning on the signal to noise ratio (*SNR*) is:

$$\begin{aligned} SNR &= 65\text{ dB} - 32\text{ dB} \\ &= 33\text{ dB} \end{aligned} \qquad (7\text{-}56)$$

and with the air conditioning off:

$$\begin{aligned} SNR &= 65\text{ dB} - 16\text{ dB} \\ &= 49\text{ dB} \end{aligned}$$

For a harmonic to be equal to −33 dB, its percentage would be:

$$100 \times 10^{\frac{-33}{20}} = 2.24\%$$

For a harmonic to be equal to −49 dB, its percentage would be:

$$100 \times 10^{\frac{-49}{20}} = 0.355\%$$

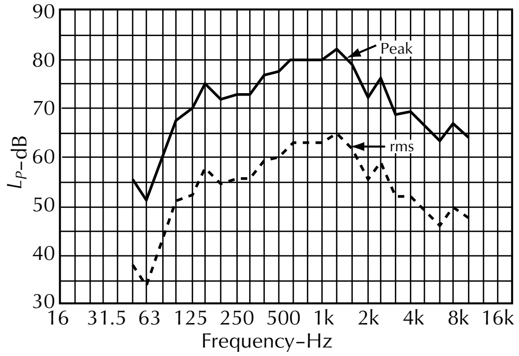


Figure 7-21. Male speech, normal level 2 feet from the microphone.

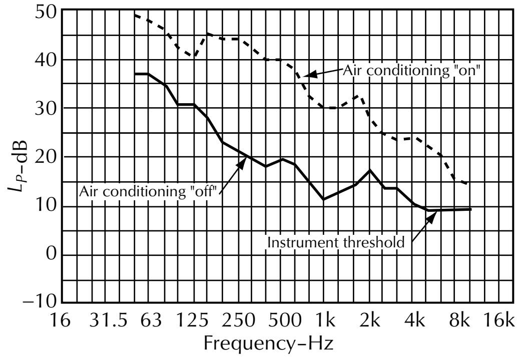


Figure 7-22. Ambient noise levels.

7.27 Decibels and Percentages

The comparison of data in decibels often needs to be expressed as percentages. The measurement of *THD* compares the harmonics with the fundamental. After finding out how many dB down each harmonic is compared to the fundamental, sum up all the harmonics and then compare their sum to the fundamental value. The difference is expressed as a percentage. The efficiency of a loudspeaker in converting electrical energy to acoustic energy is also expressed as a percentage. We know that:

$$\begin{aligned} 20\log 10 &= 20\text{ dB} \\ 20\log 100 &= 40\text{ dB} \\ 20\log 1000 &= 60\text{ dB} \end{aligned}$$

Therefore, a signal of −20 dB is 1/10 of the fundamental, or 100 × 1/10 = 10%. A signal of −40 dB is 1/100 of the fundamental, or 100 × 1/100 = 1%. A signal of −60 dB is 1/1000 of the fundamental, or 100 × 1/1000 = 0.1%. We can now turn this into an equation for finding the percentage when the level difference in decibels is known. For such ratios as voltage, SPL, and distance:

$$Percentage = 100 \times 10^{\frac{\pm dB}{20}} \qquad (7\text{-}57)$$

For power ratios:

$$Percentage = 100 \times 10^{\frac{\pm dB}{10}} \qquad (7\text{-}58)$$

Occasionally we are presented with two percentages and need the decibel difference between them. For example, two loudspeakers of otherwise identical specifications have differing efficiencies: One is 0.1% efficient, and the other is 25% efficient. If the same wattage is fed to both loudspeakers, what will be the difference in level between them in dB?

Since we are now talking about efficiency, we are talking about power ratios, not voltage ratios. We know that:

$$\begin{aligned} 10\log 10 &= 10\text{ dB} \\ 10\log 100 &= 20\text{ dB} \\ 10\log 1000 &= 30\text{ dB} \end{aligned}$$

and so forth.

A 0.1% efficiency is a power ratio of 1000 to 1, or −30 dB. We also know that −3 dB is 50% of a signal, so −6 dB would be 25%; (−6) − (−30) = 24 dB. In other words, there would be a 24 dB difference in level between these two loudspeakers when fed by the same signal. Some consumer market loudspeakers vary this much in efficiency.

7.28 Summary

The decibel is the product of the greatest engineering minds in communications early in the last century. When it is combined with the work of Oliver Heaviside and others on impedance at the turn of the 20th century, we are equipped to handle audio levels. The concepts of dB, *Z*, and dBm are the tools of the professional as well as their language.

Bibliography

V. M. Albers. *The World of Sound*. New York: Barnes, 1970.

F. Jay, Ed. *IEEE Standard Dictionary of Electrical and Electronics Terms*, 2nd ed. New York: The Institute of Electrical and Electronics Engineers, 1977.

D. N. Keast. *Measurement in Mechanical Dynamics*. New York: McGraw-Hill, 1967.

O. Read. *The Recording and Reproduction of Sound*. Indianapolis, Ind.: Howard W. Sams, 1952.

Research Council of the Academy of Motion Picture Arts and Sciences, *Motion Picture Sound Engineering*. New York: Van Nostrand, 1938.

A. Wood. *The Physics of Music*. New York: Dover, 1966.

Chapter 8

Interfacing Electrical and Acoustic Systems

by Don Davis and Eugene Patronis, Jr.

8.1 Alternating Current Circuits 113
8.2 Impedance 115
8.3 Electric Power 117
8.4 Properties of the LCR Circuit 120
8.5 Filters 121
  Octave Bandpass Filter 121
  Low Pass Filter 122
  High Pass Filter 123
  Parallel Circuits 123
  Circuit Models for Physical Inductors, Capacitors, and Resistors 125
8.6 Impedance Bridge 126
8.7 Constant Resistance Networks 128
8.8 Impedance Properties of Moving Coil Loudspeakers 129
8.9 Network Theorems 132
8.10 The Technician's Viewpoint 135
8.11 Impedance Defined 135
  Making Reactance Visible 136
  Impedance Notation 136
  Amplifier and Loudspeaker Impedances 137
  Complex Impedance 137
8.12 Handling the Acoustic Input and Output of the System 137
  The EIA Microphone Rating 139
  The Mixer Output 140
  Available Input Power 140
  Open and Matched Circuits 140
  When to Measure Z 141
  System Problems Located by Z Measurements 141
  Gain and Loss Blocks 141
  Typical Mixer Amplifier 142
8.13 Total Electrical Gain of a System 142
  A More Complicated System 142
8.14 Interfacing the Electrical Output Power to the Acoustic Environment 143
  Simplified Efficiency Calculations 143
  Damping factor 143
  Electrical Gain of a Typical System 144
  Treating Equalizer Loss 144
  Example of Gains and Losses 144
  Power Amplifier 145
8.15 Gain Structure Revisited 145
  Two Port Devices 146
  Field Adjustment by Voltage Only Analysis 146
  Voltage from Input to Output 148
  Thermal Noise Levels 149
  Microphone and Loudspeaker Polarity 149
  Microphone Interconnections 149

8.16 Conclusion 150
Bibliography 150

8.1 Alternating Current Circuits

If it were not for fear of being plagiaristic, this chapter might well be entitled *Alice in Wonderland*, with sincere apologies to Lewis Carroll. This thought stems from the puzzled expressions observed appearing on the faces of countless students when they first encounter this subject matter. This puzzlement follows from the fact that most students know quite a bit about steady state dc circuits and their mindset is to try to push the hearsay knowledge they have of ac circuits into this same framework. This does not work well at all and is similar to working on an automobile with a set of English wrenches when all of the nuts and bolts are metric. A few wrenches will fit and then only approximately. Any real craftsman knows that in order to do jobs properly, one must have the appropriate tools. If your knowledge of this subject matter is only cursory, it might be well to put aside what you may already know and begin acquiring a new set of tools or concepts as they are introduced in the following. The goal is to arrive at a thorough basic understanding rather than try to memorize a set of mysterious rules. When you thoroughly understand something, you are able to write the rulebook yourself.

The first circuit to be considered appears as Fig. 8-1. The British physicist who first analyzed this circuit is one of the boyhood heroes of many budding physics students. His name was William Thomson, later Lord Kelvin. Thomson was Professor of Physics at Glasgow University from 1846 to 1899. The analysis of this circuit, in about 1850, was perhaps the least of his achievements but was a crucial step toward making possible modern communications.

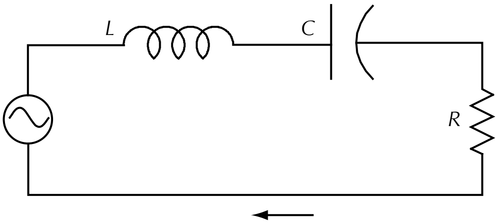


Figure 8-1. Series *LCR* circuit.

The circuit of Fig. 8-1 is a series connection of the three idealized passive circuit elements. It consists of a pure resistor, a pure capacitor, and a pure inductor all connected in series (departure from ideal behavior exhibited by real resistors, capacitors, and inductors will be discussed after development of the necessary tools). In a series circuit, the electrical current is taken as a reference as it is common to each of the circuit elements. The arrow in the diagram indicates the sense of the current when it is considered to be a positive current. This circuit, as is true of most circuits, exhibits two types of behavior. These are referred to as the transient solution and the steady state solution. For the moment, only the steady state solution will be considered. As the name implies, the steady state solution describes the behavior of the circuit after it has been connected for a reasonable time. What constitutes a reasonable time can only be answered after a study of the transient solution which will be treated in Chapter 22 *Signal Processing* on Laplace transforms.

Alternating currents are currents which either periodically or aperiodically reverse sense, i.e., are sometimes positive and sometimes negative. Periodic currents alternate between positive and negative and back again after the elapse of a definite interval of time or period, T. Even periodic alternating currents exist in many waveforms. The waveform is the form of the current as viewed on an oscilloscope or plotted on a piece of graph paper. Some common waveforms are sine, cosine, sawtooth, triangle, square, etc. Of these, the sine and cosine are essentially the same in that they differ only in the starting point reference on the oscilloscope sweep. Additionally, the sine and cosine, unlike the others, contain only a single frequency component, f. This single frequency, f, is the reciprocal of the period T. The other waveforms, even though they each have a definite period, are made up of a fundamental frequency component, which is the reciprocal of the period, as well as harmonic frequency components which are multiples of the fundamental frequency. This understanding is the result of the work of the French mathematician and physicist Jean Fourier (1768-1830). Fourier's work will be studied in more detail later.

The analysis of the circuit of Fig. 8-1 is begun by assuming the circuit current is given by

$$i = I_m \cos(2\pi f t) \tag{8-1}$$

In this statement the dependent variable is the electrical current which is denoted by the symbol i. The current is called the dependent variable because its value at any instant depends on the value which is assigned to the independent variable which is the time, t. The value of i is linked to the value of t through the functional properties of the cosine. The cosine function, of course, is tabulated in terms of an angle such as θ. In this instance the angle θ is given by

$$\theta = 2\pi f t \tag{8-2}$$

The angle θ is called the phase angle of the current. It should be noted that the value of the phase angle is directly proportional to the value of t. Starting at $t = 0$, the time which must elapse before θ takes on the value 2π is called the period, T. That is,

$$2\pi = 2\pi f T \tag{8-3}$$

or

$$f = \frac{1}{T} \tag{8-4}$$

This last equation says that the frequency is the reciprocal of the period. Finally, one defines a quantity called the angular frequency or radian frequency as follows:

$$\omega \equiv 2\pi f \tag{8-5}$$

The current can now be expressed as

$$i = I_m \cos(\omega t) \tag{8-6}$$

The maximum value that the cosine function attains is one and hence the maximum value that the current can take on is I_m. I_m is called the amplitude of the current. The current at any instant as well as the current amplitude is measured in amperes, A. An ampere is a coulomb of charge per second. Elapsed time and the period are measured in seconds, the frequency is measured in reciprocal seconds or hertz (Hz), and the angular frequency is measured in radians per second (rad/s). Fig. 8-2 is a graph of the assumed current over two periods where t is expressed in units of the period T.

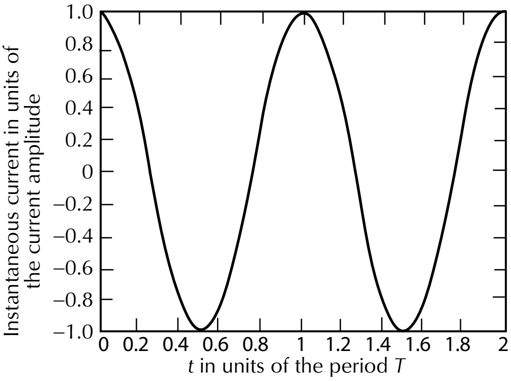


Figure 8-2. The circuit current over an interval of two periods.

The question to be answered is "What voltage must be applied across the input terminals such that the current in the circuit will be the assumed current?" This question is answered by the application of Kirchhoff's laws I and II. Law I is based on the conservation of electric charge and for a series circuit requires that the current everywhere be the same. Law II is based on the conservation of energy and requires that the voltage applied across the input terminals be equal to the sum of the voltages across the individual circuit elements in the series connected circuit. Therefore it is necessary only to determine the voltage which must exist across each circuit element and then add them up.

Unlike Lord Kelvin, who pursued the analysis by means of the differential and integral calculus, use will be made of the tools provided by Steinmetz when working with phasors. In what follows, the current will be represented by a phasor where it is understood that only the real part of the phasor represents the actual current.

$$i = I_m e^{j\omega t} \tag{8-7}$$

One can easily determine the voltage that must exist across the resistance. By the definition of resistance, the voltage across a resistance is the product of the resistance and the current. Therefore,

$$v_R = Ri = RI_m e^{j\omega t} \tag{8-8}$$

That was easy enough! The phasor representation of the voltage across the resistance is the same as the current when scaled by a factor equal to the resistance. In particular, it should be noted that the phase angle of this voltage phasor is the same as that of the current. Therefore the voltage across a resistance is in phase with the current.

When one considers the voltage across the inductance, however, the going gets a little tougher. The definition of self-inductance requires that the product of the inductance with the slope of the current curve gives the voltage across an inductance at any instant, that is,

$$v_L = L\frac{di}{dt} \tag{8-9}$$

The last equation introduces the mathematical operator d/dt. This operator is called the derivative with respect to time. In this instance, the operation is to be applied to the function representing the current and tells one to generate a new function whose value at any instant is the same as the slope of the current curve at that instant. In the language of phasors, this is accomplished simply by multiplying the current phasor by $j\omega$. Hence,

$$\begin{aligned} v_L &= j\omega L I_m e^{j\omega t} \\ &= e^{j\frac{\pi}{2}}\omega L I_m e^{j\omega t} \\ &= \omega L I_m e^{j\left(\omega t+\frac{\pi}{2}\right)} \end{aligned} \tag{8-10}$$

At this point it is worth noticing that whereas the phase angle of the circuit current is ωt, the phase angle of the voltage across the inductance is $\omega t + (\pi/2)$. The phase angle of the voltage across the inductance is greater than the phase angle of the current by an amount of $\pi/2$. This is the reason for the expression; "The voltage across an inductance leads the current."

Now for the capacitor voltage, one again goes back to fundamentals. The capacitance is defined to be the transferred charge divided by the resulting potential difference or voltage. The charge and the current are related by

$$q = \int i dt \tag{8-11}$$

Therefore, the voltage across the capacitor is given by

$$v_C = \frac{1}{C}\int i dt \tag{8-12}$$

In the language of phasors, integration with respect to time is accomplished simply by division by $j\omega$. Therefore,

$$\begin{aligned} v_C &= \frac{1}{j\omega C} I_m e^{j\omega t} \\ &= e^{-j\frac{\pi}{2}}\frac{1}{\omega C} I_m e^{j\omega t} \\ &= \frac{1}{\omega C} I_m e^{j\left(\omega t-\frac{\pi}{2}\right)} \end{aligned} \tag{8-13}$$

Note that the phase angle of the voltage across the capacitance is $\omega t - (\pi/2)$ and is less than the phase angle of the current by an amount of $\pi/2$. This is the origin of the statement; "The voltage across a capacitor lags the current."

The voltage that must be applied across the input terminals of this circuit in order to produce the assumed current is then $v_{in} = v_R + v_L + v_C$.

Fig. 8-3 depicts the phasor sum of the individual voltages. The figure is drawn for the instant of time such that $t = 0$. In the construction of the figure, it is arbitrarily assumed that the voltage across the inductance is greater than the voltage across the capacitance.

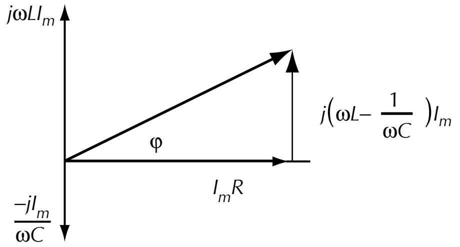


Figure 8-3. Phasor sum of resistor, capacitor, and inductor voltages.

An examination of Fig. 8-3 will indicate that the phasor sum of the individual voltages yields

$$v_{in} = I_m\left[R + j\left(\omega L - \frac{1}{\omega C}\right)\right]e^{j\omega t} \tag{8-14}$$

8.2 Impedance

All of the foregoing analysis can be made much more compact by defining a new term called the circuit impedance, Z. The circuit impedance, Z, is defined to be the ratio of the complex applied voltage to the complex current which results from the application of that voltage.

$$Z \equiv \frac{v_{in}}{i} \tag{8-15}$$

Upon applying this definition, the impedance of the series LCR is found to be

$$\begin{aligned} Z &= \frac{v_{in}}{i} \\ &= \frac{I_m\left[R + j\left(\omega L - \frac{1}{\omega C}\right)\right]e^{j\omega t}}{I_m e^{j\omega t}} \\ &= \left[R + j\left(\omega L - \frac{1}{\omega C}\right)\right] \end{aligned} \tag{8-16}$$

The last expression is the complex circuit impedance written in rectangular form. When written in the complex exponential form it appears as

$$Z = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}\, e^{j\varphi} \tag{8-17}$$

with the angle φ being given by

$$\varphi = \operatorname{atan}\frac{\left(\omega L - \frac{1}{\omega C}\right)}{R} \tag{8-18}$$

In general, the definition of the impedance implies a knowledge of two quantities. One is called the magnitude of the impedance, which is the ratio of the applied voltage amplitude to the amplitude of the resulting current. The other is called the angle of the impedance, which is the phase angle of the applied voltage minus the phase angle of the current that results from the application of that voltage. The magnitude of the impedance is denoted by $|Z|$. For the case at hand,

$$\begin{aligned} |Z| &= \frac{V_m}{I_m} \\ &= \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2} \end{aligned} \tag{8-19}$$

where,
V_m is the amplitude of the applied voltage,
I_m is the amplitude of the circuit current.

At this point, it is well to re-examine the circuit in a slightly different way. The circuit consists of a pure inductor, a pure capacitor, and a pure resistor connected in series. Each of these circuit elements has its own impedance. These respective impedances are

$$Z_L = j\omega L \tag{8-20}$$

$$Z_C = \frac{-j}{\omega C} \tag{8-21}$$

$$Z_R = R \tag{8-22}$$

The impedance of the inductance is purely imaginary. Such an impedance is termed a reactance. In this instance it is a positive reactance as it falls on the positive imaginary axis. This reactance is denoted by X_L with

$$X_L = \omega L \tag{8-23}$$

The impedance of the capacitance is also purely imaginary indicating that it is also a reactance. In this case, however, it is a negative reactance as it falls on the negative imaginary axis. This reactance is denoted by X_C with

$$X_C = \frac{-1}{\omega C} \tag{8-24}$$

Finally, the impedance of the resistance is purely real with

$$Z_R = R \tag{8-25}$$

The total impedance of the circuit is thus

$$\begin{aligned} Z &= Z_L + Z_C + Z_R \\ &= R + j\left(\omega L - \frac{1}{\omega C}\right) \end{aligned} \tag{8-26}$$

when written in rectangular form, or

$$Z = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}\, e^{j\,\mathrm{atan}\left(\frac{\omega L - \frac{1}{\omega C}}{R}\right)} \tag{8-27}$$

when written in exponential form. Notice that the magnitude of the total impedance is the square root of the sum of the resistance squared and the net reactance squared while the angle of the impedance is the angle whose tangent is the ratio of the net reactance to the resistance.

Impedances may be summed as was done above or they may be summed graphically by drawing a diagram in the complex plane. This diagram is similar to a phasor diagram as they are both drawn in the complex plane. Unlike the phasor diagram, the impedance diagram does not rotate with the passage of time. An impedance diagram is fixed, not changing with time as long as the circuit components remain unchanged. Fig. 8-4 is an impedance diagram for the present circuit.

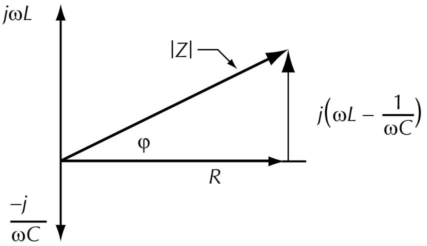


Figure 8-4. Impedance diagram for the series *LCR* circuit.

Now that the concept of impedance has been introduced and, hopefully, is understood, circuit analysis is greatly simplified. The procedure to be employed is outlined in the following steps assuming that the current is a known quantity.

1. Calculate the total circuit impedance and express it in exponential form.
2. Express current as a phasor.
3. Multiply current phasor by the total impedance to obtain the voltage phasor.
4. Take the real part of the voltage phasor to obtain the actual voltage as a function of time.

When the applied voltage is a known quantity, the procedure is listed in the following steps:

1. Calculate the total circuit impedance and express it in exponential form.
2. Express the applied voltage as a phasor.
3. Divide the voltage phasor by the circuit impedance to obtain the current phasor.
4. Take the real part of the current phasor to obtain the actual current as a function of time.

All of this is best illustrated through a numerical example.

A series *LCR* circuit supports a current

$$i = 3\ \mathrm{A}\cos\left(\frac{500}{\mathrm{s}}t\right)$$

The series inductance is 0.04 henry (H), the series capacitance is 0.0002 farad (F), and the series resistance is 20 ohms (Ω). Find the voltage (V) which must be applied to produce the given current (I).

The impedance of the resistor is 20 Ω

The impedance of the inductance is

$$\begin{aligned} j\omega L &= j\left(\frac{500}{\mathrm{s}}\right)(0.04\ \mathrm{H}) \\ &= j20\ \Omega \end{aligned}$$

The impedance of the capacitor is

$$\begin{aligned} \frac{-j}{\omega C} &= \frac{-j}{\left(\frac{500}{\mathrm{s}}\right)(0.0002\ \mathrm{F})} \\ &= -j10\ \Omega \end{aligned}$$

The total impedance is

$$\begin{aligned} Z &= 20 + j(20-10) \\ &= \sqrt{20^2 + (20-10)^2}\,e^{j\,\mathrm{atan}\left(\frac{20-10}{20}\right)} \\ &= 22.361\,e^{j0.464}\ \Omega \end{aligned}$$

The current phasor is

$$i = 3\ \mathrm{A}e^{j\left(\frac{500}{\mathrm{s}}t\right)}$$

The voltage phasor is

$$\begin{aligned} v &= iZ \\ &= 67.083\mathrm{V}e^{j\left(\frac{500}{\mathrm{s}}t + 0.464\right)} \end{aligned}$$

The actual voltage is thus

$$v = 67.083\ \mathrm{V}\ \cos\left(\frac{500}{\mathrm{s}}t + 0.464\right)$$

Alternately let the voltage applied to the same circuit be

$$v = 25\ \mathrm{V}\ \cos\left(\frac{500}{\mathrm{s}}t\right)$$

What is the circuit current in this case?

The total impedance is

$$Z = 22.361\,e^{j0.464}\Omega$$

The voltage phasor is

$$v = 25\ \mathrm{V}\ e^{j\left(\frac{500}{\mathrm{s}}t\right)}$$

The current phasor is

$$i = \frac{v}{Z} = 1.118\ \mathrm{A}\ e^{j\left(\frac{500}{\mathrm{s}}t - 0.464\right)}$$

The actual current is thus

$$i = 1.118\ \mathrm{A}\ \cos\left(\frac{500}{\mathrm{s}}t - 0.464\right)$$

8.3 Electric Power

Power is the rate of doing work. In this regard, one considers both instantaneous power and average power. Instantaneous power is the slope of the work curve versus time and thus is the work done in a vanishingly small interval of time divided by the small interval of time. The average power over an appreciable interval of time is the total work done in the interval of time divided by the length of the time interval. In electrical systems, the voltage represents electrical energy per unit of electric charge while the current represents charge per unit time. The instantaneous electric power is the product of the instantaneous applied voltage at a given point in time with the instantaneous current at that same point in time.

Consider a pure resistance in which there exists a sinusoidal current. From our previous work, it is possible to write the following relationships.

$$i = I_m \cos\left(\frac{2\pi t}{T}\right) \tag{8-28}$$

$$v_R = V_m \cos\left(\frac{2\pi t}{T}\right) \tag{8-29}$$

$$V_m = I_m R \tag{8-30}$$

$$\begin{aligned} p &= i v_R \\ &= \frac{V_m^{\ 2}}{R} \cos^2\left(\frac{2\pi t}{T}\right) \end{aligned} \tag{8-31}$$

The instantaneous power is symbolized by p and is displayed in Fig. 8-5 for an interval equal to one period, T. The area beneath the curve is colored. This area is equal to the electrical energy converted into heat in the resistance in the time T.

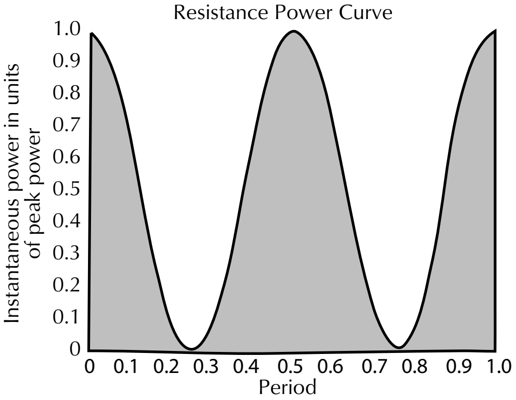


Figure 8-5. Instantaneous power developed in a resistance by a sinusoidal current.

The total work done by the applied alternating voltage in the time T is denoted by W with W being given by

$$W = 0.5\left(\frac{V_m^{\ 2}}{R}\right)T \tag{8-32}$$

The average power in the resistance is the total work done divided by the time spent in performing the work, hence the average power in the resistance, denoted by P, appears as

$$\begin{aligned} P &= \frac{W}{T} \\ &= 0.5\left(\frac{V_m^{\ 2}}{R}\right) \end{aligned} \tag{8-33}$$

At this juncture it is reasonable to inquire, "What constant dc voltage applied to the resistance would produce the same average power as does the sinusoidal alternating voltage?" As is well known, the power developed in a resistance by a steady dc voltage V is the voltage squared divided by the resistance. The posed question may be answered by solving the equation

$$0.5\left(\frac{V_m^{\ 2}}{R}\right) = \frac{V^2}{R}$$

from which

$$V = \frac{V_m}{\sqrt{2}} \tag{8-34}$$

This is called the effective or root mean square value of the sinusoidal alternating voltage. This same analysis could have been done in terms of the current rather than the voltage with the result that the effective or root mean square current would have been found to be

$$I = \frac{I_m}{\sqrt{2}} \tag{8-35}$$

A word of caution, the numerical factor of square root of two in these equations is appropriate for a sinusoidal voltage or current. In general, however, other wave shapes such as triangle, square, etc. will have different numerical values associated with their root mean square values.

Turn now to a pure inductance in which there exists a sinusoidal current. Again from the previous work, it is possible to write

$$i = I_m \cos\left(\frac{2\pi t}{T}\right) \tag{8-36}$$

$$v_L = V_m \cos\left(\frac{2\pi t}{T} + \frac{\pi}{2}\right) \tag{8-37}$$

$$V_m = I_m \omega L \tag{8-38}$$

$$\begin{aligned} p &= i v_L \\ &= \frac{V_m^{\ 2}}{\omega L} \cos\left(\frac{2\pi t}{T}\right) \cos\left(\frac{2\pi t}{T} + \frac{\pi}{2}\right) \end{aligned} \tag{8-39}$$

The instantaneous power curve for the inductance is displayed in Fig. 8-6. With regard to Fig. 8-6, the function being plotted is

$$p = iv_L$$
$$= \frac{V_m^{\ 2}}{\omega L}\cos\left(\frac{2\pi}{T}\right)\cos\left(\frac{2\pi}{T}+\frac{\pi}{2}\right) \qquad (8\text{-}40)$$

where,

$$\frac{V_m^{\ 2}}{\omega L} = V_m I_m$$

Furthermore, the product of the two cosine functions has a maximum value of 0.5 so the maximum ordinate on the graph is 0.5 when the vertical axis is labeled in units of $V_m I_m$.

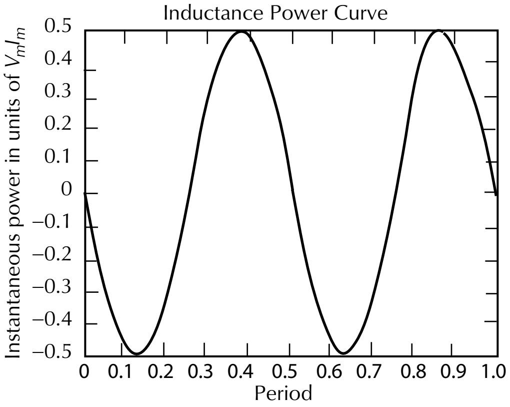


Figure 8-6. Instantaneous power developed in an inductance by a sinusoidal current.

Note that in contrast to the resistive case where the power curve was always positive, here the power curve is alternately negative and positive. Furthermore, when p is negative, the area bounded by the curve is also negative. When p is positive, the area bounded by the curve is also positive. The interesting point here is that over an interval of one period or any integral number of periods, the net area and hence the total work done is zero. This means that the average electrical power in a pure inductance and, indeed, in any reactance is zero. When the power curve is positive, the inductance is receiving energy from the external circuit and is storing this energy in the magnetic field which surrounds the inductance. When the power curve is negative, the inductance's stored magnetic energy is being converted back into electrical energy which is returned to the external circuit. A similar phenomenon occurs with a capacitive reactance except here the storage medium is the electric field in the capacitance. In summary, the average power in any reactance, positive or negative, is zero. This means that the average power dissipated in any impedance is associated only with the real or resistive part of the impedance.

In conclusion, consider a general impedance in which there exists a sinusoidal current. The appropriate equations are

$$i = I_m\cos\left(\frac{2\pi t}{T}\right) \qquad (8\text{-}41)$$

$$v = I_m|Z|\cos\left(\frac{2\pi t}{T}+\varphi\right) \qquad (8\text{-}42)$$

$$p = I_m^{\ 2}|Z|\cos\left(\frac{2\pi t}{T}\right)\cos\left(\frac{2\pi t}{T}+\varphi\right) \qquad (8\text{-}43)$$

Fig. 8-7 is the instantaneous power curve for this general case where an arbitrary positive value has been assigned to the angle of the impedance.

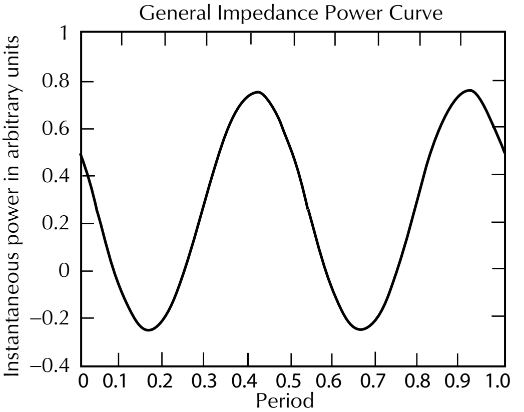


Figure 8-7. Instantaneous power in a general impedance.

From Fig. 8-7 it is apparent that the instantaneous power has both positive and negative excursions with the positive excursions being larger than the negative ones. The positive areas bounded by the curve are larger than the negative areas and hence, there is an average power greater than zero. The average power in this case is given by

$$P = 0.5I_m^{\ 2}|Z|\cos\varphi \qquad (8\text{-}44)$$

The quantity cosine of the angle of the impedance is called the power factor. The angle φ falls in the interval ±π/2. The cosine is always positive in this interval, therefore the average power is equal to or greater than zero. The multiplication of the magnitude of the impedance by the cosine of the angle of the impedance has the effect of selecting out just the real part of the general impedance and hence the equation could just as well be written

$$P = 0.5 I_m^2 R = I^2 R \tag{8-45}$$

8.4 Properties of the *LCR* Circuit

It is now possible to explore the very useful properties of the series *LCR* circuit. This is facilitated by an examination of Fig. 8-8 which is a modification of Fig. 8-1.

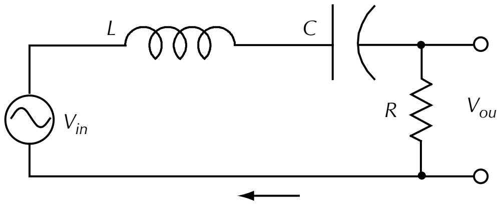


Figure 8-8. A revised drawing of the *LCR* circuit.

Consider the generator to be an ideal variable frequency voltage source in which the frequency can take on any value between zero and infinity. The subject to be studied is the relationship which exists in the steady state between the output voltage, which is the voltage across the resistor, and the input voltage, which is the voltage applied to the entire circuit. Therefore,

$$\begin{aligned}\frac{V_{out}}{V_{in}} &= \frac{iR}{iZ}\\ &= \frac{R}{R + j\left(\omega L - \frac{1}{\omega C}\right)}\end{aligned} \tag{8-46}$$

The objective at this point is to extract as much information as is possible from this equation. The first observation, of course, is that the relationship is complex which implies both magnitude as well as phase information. The magnitude information will be explored first. The magnitude of the ratio of the voltage out to the voltage in is given by

$$\frac{R}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}} \tag{8-47}$$

This expression takes on its largest value when the denominator is the least. The denominator has its least value when the term in the parentheses is zero. There is one positive value of ω for which this is true. Call this value ω_0. This value is found through the following steps.

$$\begin{aligned}\omega_0 L - \frac{1}{\omega_0 C} &= 0\\ \omega_0^2 L &= \frac{1}{C}\\ \omega_0 &= \sqrt{\frac{1}{LC}}\end{aligned} \tag{8-48}$$

When ω is equal to ω_0, the ratio is 1 and the size of the output voltage is the same as that of the input voltage. For any other value of the angular frequency, the parentheses quantity is not zero and the ratio is less than one. This unique value of the angular frequency is called the resonant angular frequency. The resonant frequency itself is f_0 which, of course, is ω_0 divided by 2π. The output voltage, which is the voltage across the resistor, is at a maximum at resonance. The output power, considered to be the power in the resistance, is also at a maximum at resonance and is equal both instantaneously and on the average to the power supplied by the generator.

There are two other angular frequencies of particular interest. One of these is higher than ω_0 while the other is lower. Denote these by ω_H and ω_L, respectively. At these values of the angular frequency, the voltage ratio magnitude is reduced to $1/\sqrt{2}$. For these values of the angular frequency, the power, which depends on the voltage squared, is one-half of its maximum value. These values of the angular frequency constitute the half-power points. At these points, it must be true that

$$\left(\omega L - \frac{1}{\omega C}\right)^2 = R^2 \tag{8-49}$$

This leads to two other quadratic equations

$$\omega_H^2 L - \omega_H R - \frac{1}{C} = 0 \tag{8-50}$$

$$\omega_L^2 L + \omega_L R - \frac{1}{C} = 0 \tag{8-51}$$

Solving these equations leads to the result

$$\omega_H = \sqrt{\frac{R^2}{4L^2} + \omega_0^2} + \frac{R}{2L} \tag{8-52}$$

$$\omega_L = \sqrt{\frac{R^2}{4L^2} + \omega_0^2} - \frac{R}{2L} \tag{8-53}$$

It is instructive to examine the width of the resonance represented by the difference between the higher and lower half-power point angular frequencies. This difference is

$$\omega_H - \omega_L = \frac{R}{L} \tag{8-54}$$

Whether the resonance is well defined or sharp or whether it is poorly defined or broad is gauged by a quantity called the quality factor or Q. Q is defined in such a way that it is a large number for a sharp, well-defined resonance and is small for a poorly defined or broad resonance.

$$Q \equiv \frac{\omega_0}{\omega_H - \omega_L} \tag{8-55}$$

Substitution of the known width of the resonance leads to

$$\begin{aligned} Q &= \frac{\omega_0 L}{R} \\ &= \frac{\sqrt{\frac{L}{C}}}{R} \end{aligned} \tag{8-56}$$

A few final observations before leaving the subject of the magnitude behavior of the voltage ratio. An inspection of the magnitude expression reveals that the voltage ratio is zero at the angular frequency extremes of zero and infinity. Additionally, after some algebra, one finds that

$$\omega_0 = \sqrt{\omega_H \omega_L} \tag{8-57}$$

This last relationship describes the fact that the half power points are disposed about the maximum power point with geometric rather than arithmetic symmetry. This will be readily apparent after graphing the voltage ratio function.

In order to explore the phase behavior of the ratio of the output voltage to the input voltage, it is necessary to express this ratio in the exponential rather than the rectangular form.

$$\frac{V_{out}}{V_{in}} = \frac{R}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}} e^{j\left[-\mathrm{atan}\left(\frac{\omega L - \frac{1}{\omega C}}{R}\right)\right]} \tag{8-58}$$

Inspection of the equation indicates that φ, which in this case is the phase angle of the output voltage less the phase angle of the input voltage, appears as

$$\varphi = -\mathrm{atan}\frac{\left(\omega L - \frac{1}{\omega C}\right)}{R} \tag{8-59}$$

When ω is zero, φ is $+\pi/2$. When ω approaches infinity, φ approaches $-\pi/2$. When $\omega = \omega_0$, $\varphi = 0$. Finally, at the half power points, $\varphi = \pm(\pi/4)$ while being positive at the lower half power point. It is worth repeating that when $\omega = \omega_0$ not only is the magnitude of the voltage out equal to the magnitude of the voltage in, but also the phase of the voltage out is identical to the phase of the voltage in. At this particular frequency, it is just as if the series combination of the inductor and capacitor constitute a short circuit! The *LCR* circuit, when arranged in the manner of Fig. 8-8 constitutes the original bandpass filter from which all others have followed. This fact alone makes Kelvin's contribution so significant. An even better appreciation for the behavior of this filter can be had from an examination of its response curves. This will be explored in two ways by employing linear axes and combinations of logarithmic axes. In doing this, numerical values will be employed that yield a bandpass filter that might well be encountered in audio engineering.

8.5 Filters

8.5.1 Octave Bandpass Filter

The exercise is to design a simple bandpass filter which has a bandwidth of one octave, a band center frequency of 1 kHz, and a termination resistance of 600 Ω. Two things are known immediately from the statement of the problem. It is a given that $R = 600\,\Omega$ and that $f_0 = 1000\,\text{Hz}$. This requires that $\omega_0 = 2000\pi\,\text{rad/s}$. The problem then becomes one of determining the appropriate values required of L and C. As there are two unknowns, two linearly independent equations, which may be solved simultaneously, are required. Such equations have been presented in the course of the analysis. Those equations are

$$\omega_0 = \sqrt{\frac{1}{LC}} \tag{8-60}$$

$$Q = \frac{\sqrt{\frac{L}{C}}}{R} \tag{8-61}$$

ω_0 is known thus taking care of the first equation, but what is the value of Q? The value required of Q

is implicit in the statement of the problem. The bandwidth of the filter is to be one octave. Bandwidth is defined to be the frequency interval between the half power points of the filter. An octave bandpass filter is then one in which ω_H is twice as large as ω_L. Therefore by consulting the defining equation for Q and substituting the relationship between the half power points for the octave filter, Q can be extracted as

$$\begin{aligned} Q &= \frac{\omega_0}{\omega_H - \omega_L} \\ &= \frac{\sqrt{\omega_H \omega_L}}{\omega_H - \omega_L} \\ &= \frac{\sqrt{2\omega_L{}^2}}{2\omega_L - \omega_L} \\ &= \sqrt{2} \end{aligned}$$

With a value of Q in hand, separate values can be extracted for L and C. The results are

$R = 600\,\Omega$,

$L = 0.135\,\text{H}$,

$C = 0.1876\,\mu\text{F}$.

Fig. 8-9 is a plot of the behavior of the magnitude of the ratio of the output voltage to the input voltage for the filter. This is referred to as the amplitude response. In this instance, linear axes are employed and the lack of symmetry is readily apparent.

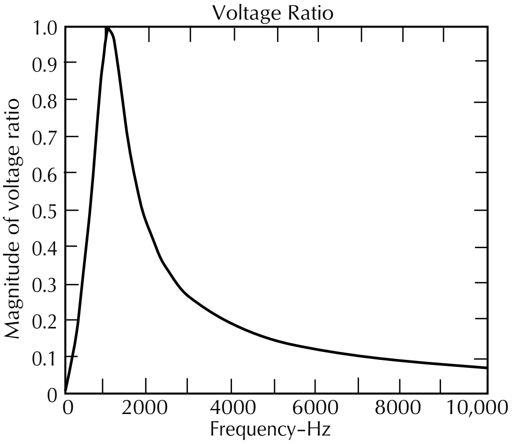


Figure 8-9. Amplitude response of filter with linear axes.

Fig. 8-10 is the amplitude response of the filter depicted with log axis for frequency and with the magnitude expressed in decibels which also constitutes a log axis.

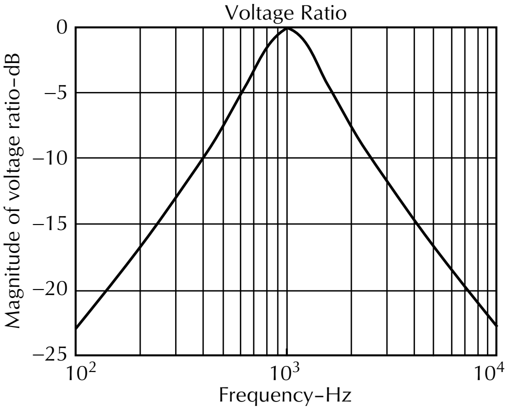


Figure 8-10. Amplitude response of the filter displayed with logarithmic axes.

In Fig. 8-10 the symmetry is obvious. The fact that the symmetry is obvious on a log axis tells one that the function being plotted is geometrically symmetric. This is just one of the reasons why response curves are usually drawn with log frequency axes.

Fig. 8-11 is the phase response of the octave bandpass filter displayed with a linear frequency axis while Fig. 8-12 displays the phase response with a logarithmic frequency axis.

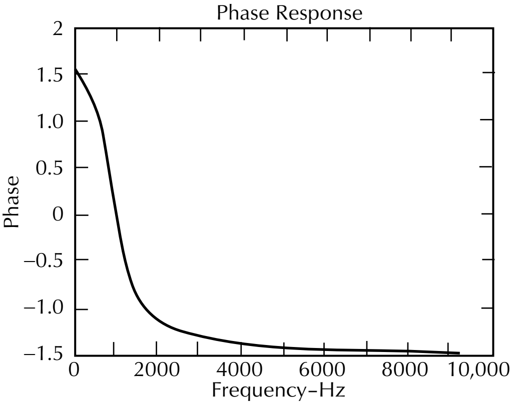


Figure 8-11. Phase response with linear frequency axis.

8.5.2 Low Pass Filter

In the equations describing the amplitude and phase response of the bandpass filter, one can readily extract the corresponding ones for a low pass filter by letting C become infinite mathematically. Physically this amounts to replacing the capacitor with a short circuit. If this appears puzzling, imagine increasing the capacitance by placing the plates

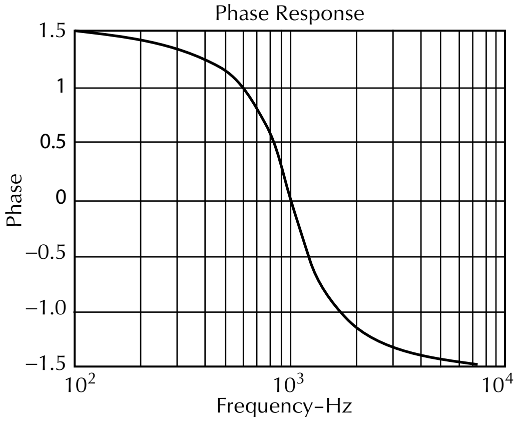


Figure 8-12. Phase response with logarithmic frequency axis.

closer and closer together until in the limit the separation is zero. The low pass equations then are

$$\left|\frac{V_{out}}{V_{in}}\right| = \frac{R}{\sqrt{R^2 + (\omega L)^2}} \tag{8-62}$$

$$\varphi = -\mathrm{atan}\left(\frac{\omega L}{R}\right) \tag{8-63}$$

These responses are displayed in Figs. 8-13 and 8-14 employing the previous values for L and R.

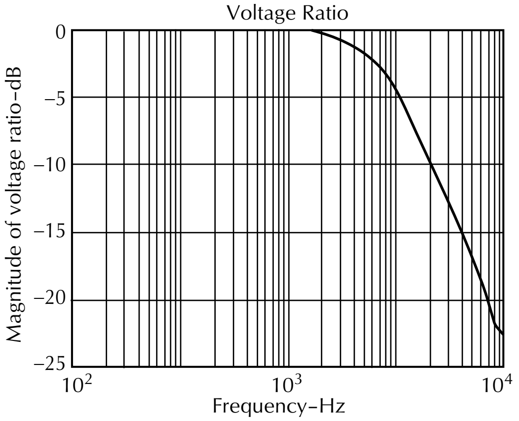


Figure 8-13. Amplitude response of low pass filter.

8.5.3 High Pass Filter

The general equations for the bandpass filter can similarly be converted to those for a high pass filter by letting L become equal to zero. In this instance, the amplitude and phase response equations become

$$\left|\frac{V_{out}}{V_{in}}\right| = \frac{R}{\sqrt{R^2 + \left(\frac{1}{\omega C}\right)^2}} \tag{8-64}$$

$$\varphi = -\mathrm{atan}\left(\frac{-1}{R\omega C}\right) \tag{8-65}$$

The results appear in Figs. 8-15 and 8-16 while again using the given values of R and C.

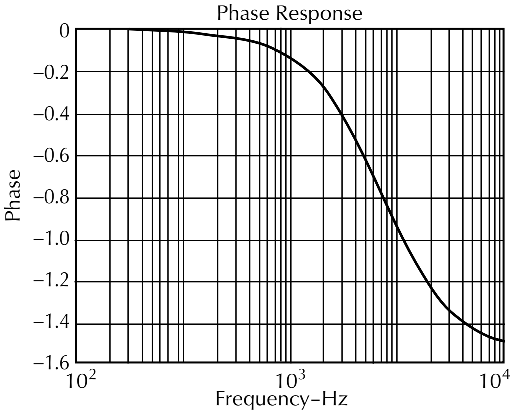


Figure 8-14. Phase response of low pass filter.

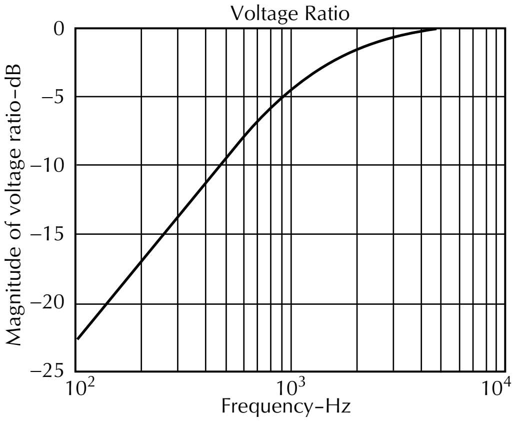


Figure 8-15. Amplitude response of high pass filter.

8.5.4 Parallel Circuits

In the course of an ordinary day it is not uncommon to encounter a sign which proclaims NO ADMITTANCE. This could just as logically read INFINITE IMPEDANCE because admittance is defined to be the ratio of the resulting current to the applied voltage. Admittance and impedance are reciprocally related quantities.

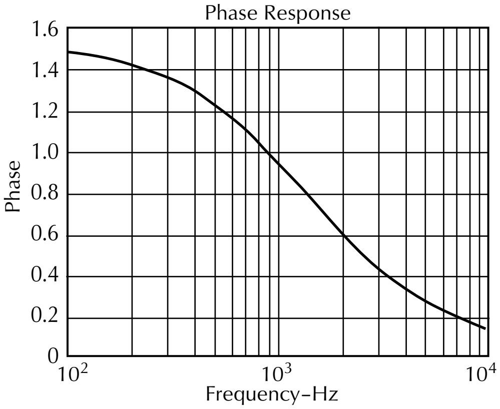


Figure 8-16. Phase response of high pass filter.

$$Y \equiv \frac{i}{v} = \frac{1}{Z} \tag{8-66}$$

In series circuits, the current is the common factor and it is necessary to sum the voltages across the individual elements. This leads to the result that the total circuit impedance is the sum of the impedances of the individual circuit elements. In a parallel circuit, the voltage is the common element and one must sum the individual currents. This leads to the result that the total circuit admittance is the sum of the admittances of the individual circuit elements which are connected in parallel.

The admittance of a pure resistor is called the conductance and is symbolized by the letter G with

$$G = \frac{1}{R} \tag{8-67}$$

G has the dimension of reciprocal ohms with a reciprocal ohm being called a siemen.

The admittance of a pure inductance is called a negative susceptance and is symbolized by the letter B with

$$B = -\frac{1}{\omega L} \tag{8-68}$$

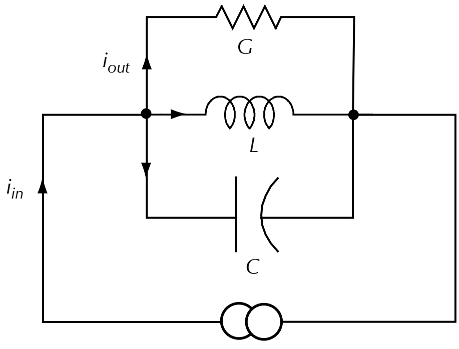
The admittance of a pure capacitance is called a positive susceptance and is symbolized by the letter B with

$$B = \omega C \tag{8-69}$$

A generalized admittance would appear then as

$$Y = G \pm jB \tag{8-70}$$

A parallel circuit of interest and of practical application consists of a pure inductance, a pure capacitance, and a pure conductance all connected.



Constant current source

Figure 8-17. This is the dual of the series *LCR* with a voltage source.

This circuit is said to be the dual of the circuit displayed in Fig. 8-8. In a dual, constant voltage sources are replaced by constant current sources. A series connection is replaced by a parallel connection. Voltages are replaced by currents and impedances are replaced by admittances. The total admittance of a parallel combination is the sum of the admittances in the individual parallel branches. The voltage across the parallel combination, from the definition of admittance, is the total current divided by the total admittance. Therefore,

$$Y = G + j\left(\omega C - \frac{1}{\omega L}\right) \tag{8-71}$$

$$v = \frac{i_{in}}{Y} = \frac{i_{in}}{G + j\left(\omega C - \frac{1}{\omega L}\right)} \tag{8-72}$$

$$i_{out} = vG = \frac{i_{in} G}{G + j\left(\omega C - \frac{1}{\omega L}\right)} \tag{8-73}$$

$$\frac{i_{out}}{i_{in}} = \frac{G}{G + j\left(\omega C - \frac{1}{\omega L}\right)} \tag{8-74}$$

This circuit also displays a resonance condition which occurs again with $\omega_0 = \sqrt{1/(LC)}$. When $\omega = \omega_0$, the net susceptance is zero. The currents in the inductance and the capacitance are of equal magnitude but of opposite polarity such that their sum is zero. All of the current furnished by the current source passes through the conductance.

The useful power in this circuit is considered to be that which appears in the conductance. This power is a maximum at resonance and is zero at the frequency extremes. The half power points for this circuit are

$$\omega_H = \frac{G}{2C} + \sqrt{\frac{G^2}{4C^2} + \omega_0^{\ 2}} \tag{8-75}$$

$$\omega_L = -\frac{G}{2C} + \sqrt{\frac{G^2}{4C^2} + \omega_0^{\ 2}} \tag{8-76}$$

Finally, the quality factor for this circuit appears as

$$Q = \frac{\sqrt{\frac{C}{L}}}{G} \tag{8-77}$$

This circuit is a workhorse for bandpass amplifiers wherein the current source is an active element such as a bipolar or field effect transistor or even a pentode vacuum tube. The parallel L and C form what is termed a tank circuit with the conductance representing the useful load which the amplifier feeds. The center of the pass band is determined by the LC combination while the width of the bandpass is governed by Q. An octave bandpass filter is readily constructed using this topology. Requiring the resonance to occur at 1 kHz with a width of one octave requires the following values for the circuit parameters:

1. $L = 0.06752\,\text{H}$.
2. $C = 0.3751\,\mu\text{F}$.
3. $G = 0.001667\,\text{S}$.

The performance of such a circuit is depicted in Fig. 8-18 which should be compared with Fig. 8-10.

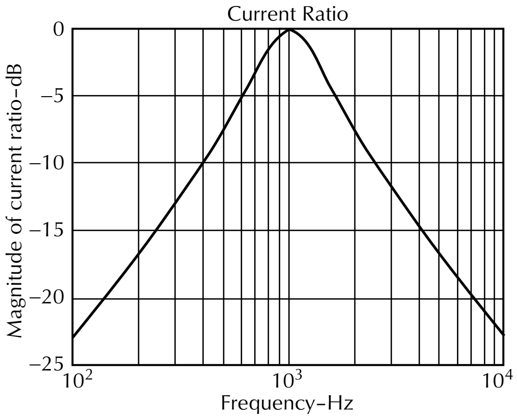


Figure 8-18. Current driven bandpass filter.

8.5.5 Circuit Models for Physical Inductors, Capacitors, and Resistors

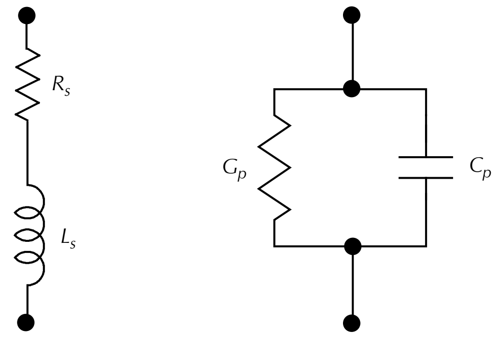
In the previous analyses, the circuit components were considered to be pure with each acting as pure inductance, capacitance or resistance, respectively. Physical inductors, capacitors, and resistors exhibit behaviors that can and do depart from the ideal. This departure from ideal behavior may or may not be significant depending upon the degree. One must at least be aware of the possibility and be familiar with tools that are capable of quantifying any departure from the ideal.

A pure inductance is an energy storage element that stores energy in the magnetic field in the space surrounding the inductance with no attendant loss of energy. This energy is often referred to as kinetic energy as the amount of this stored energy is proportional to the square of the current and the current exists because charge is in motion. A physical inductor is usually made by taking an insulated wire formed from a good conductor, such as copper, and winding this wire into a coil. Such a coil will exhibit an amount of self-inductance, which depends on the geometry of the coil. Additionally, the inductance value increases generally with the square of the number of turns used in forming the coil. Finally, the inductance value can be greatly influenced by the presence of a core of ferromagnetic material. The wire, from which the inductor is wound, and a ferromagnetic core, if present, are sources of energy loss.

A ferromagnetic core provides two loss mechanisms for time varying currents. The first of these is due to eddy currents induced in the core by the time varying magnetic field. This effect can be minimized by constructing the core from insulated laminations of the core material or by casting the core from ceramic ferrites. The second effect is magnetic hysteresis, which is an inherent property of ferromagnetism. The core losses also amount to the conversion of electrical energy into heat. The conductive losses and the core losses can together be represented by some equivalent resistance being associated with the inductor. A simple model for a physical inductor, which accounts for these effects, consists of a pure resistance in series with a pure inductance. These elements must be sized so as to account for the actual behavior of the physical inductor.

A pure capacitance is an energy storage element that stores energy in the electric field which exists in the space between the conductors forming the capacitor. This energy is called potential energy, as the amount of stored energy is proportional to the square of the amount of static charge on the conductors. Most capacitors employ dielectric materials filling the space between the conducting plates forming the capacitor. The dielectric serves as insulation between the plates. Additionally, the polarization occurring in the molecular structure of the dielectric greatly enhances the amount of capacitance obtainable.

Dielectrics, however, are not perfect insulators and attendant to molecular polarization is a relaxation phenomenon. Both of these properties lead to energy losses. For example, an isolated charged capacitor will slowly discharge over a period of time. A simple model for a capacitor then, consists of a pure capacitance in parallel with a pure conductance adequately sized to account for the actual behavior of the physical capacitor in question. In general, it is possible to construct actual capacitors, which more closely approach ideal behavior than can be the case with actual inductors. The simple models for both the physical inductor and the physical capacitor appear in Fig. 8-19.



A. Inductor B. Capacitor

Figure 8-19. Simple models for a physical inductor and physical capacitor.

Both of these models require modifications when working at high frequencies and, in particular, frequencies above the audio band. In this region, the effect of distributed capacitance between the turns of the coil forming the inductor becomes significant. This is accounted for by placing a suitably sized capacitance connected in parallel with the terminals of Fig. 8-19A.

In the case of the capacitor, in the high frequency region the leads and the conductors forming the capacitor plates exhibit some inductance of significance. This is accounted for by placing a small inductance in series with the terminals of Fig. 8-19B.

Physical resistors of the molded composition or metal film type are more nearly ideal, at least at audio frequencies. Wire wound power resistors can and do exhibit inductance unless special winding techniques are employed in their construction. Such components can be modeled as in Fig. 8-19A with the series resistance being the dominant term. Even non-inductive resistors at sufficiently high frequencies will suffer from some stray capacitance existing between the resistor terminals. The size of this capacitance depends in part on the immediate environment of the resistor. For low values of resistance, high conductance, even this effect is usually negligible. For large values of resistance, low conductance, the stray capacitance becomes significant. When this is the case, the model of Fig. 8-19B becomes appropriate.

The figure of merit, which assesses how closely a physical inductor approaches being a pure inductance in a given case, is the inductor quality factor, denoted by the symbol Q. The Q for an inductor is the ratio of the magnitude of the inductive reactance to the resistance associated with the inductor. Q then is given by

$$Q = \frac{\omega L_s}{R_s} \tag{8-78}$$

The figure of merit in the capacitive case is called the dissipation factor; denoted by the symbol D. The D for a capacitor is the ratio of the conductance to the magnitude of the capacitive susceptance. D then is given by

$$D = \frac{G_p}{\omega C_p} \tag{8-79}$$

An instrument, which facilitates the direct measurement of these properties of circuit components, is discussed in the next section.

8.6 Impedance Bridge

Fig. 8-20 depicts a generalized bridge circuit. The determination of the currents that exist in such a circuit is an interesting problem in itself. This problem will be solved in the course of gaining an understanding of the operation of an impedance bridge.

In the bridge circuit of Fig. 8-20, the voltage source is an oscillator that furnishes a constant amplitude sinusoidal voltage with a selected frequency. In circuits of this type in which the connections are not simply series and parallel combinations, the procedure is to first assign identifiable currents to each closed conducting loop or mesh. The loops are arbitrary as long as they provide for the existence of a current in each of the circuit impedances. The impedances in Fig. 8-20 are assumed to be general in nature. In a particular case, an individual element may be real, imaginary, or complex. In the case of an impedance bridge, the

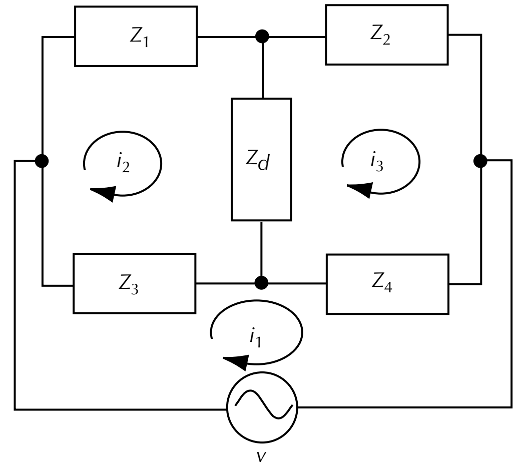


Figure 8-20. A general bridge circuit.

element Z_d represents the impedance of an alternating current indicating instrument such as a meter. For audio frequencies, this meter might be as simple as a pair of headphones.

The next step is to set up a system of linearly independent equations that will allow the determination of the unknown currents. In the case at hand there are three unknown currents and a system of three simultaneous linearly independent equations is required. These equations are obtained by applying Kirchhoff's second law to each of the chosen closed loops. These equations are then solved for the individual currents by applying the rules of algebra to the set of simultaneous equations. The set of equations applicable to the loops indicated in Fig. 8-20 are

$$(Z_3+Z_4)i_1 - Z_3 i_2 - Z_4 i_3 = v$$

$$-Z_3 i_1 + (Z_1+Z_3+Z_d)i_2 - Z_d i_3 = 0$$

$$-Z_4 i_1 - Z_d i_2 + (Z_2+Z_4+Z_d)i_3 = 0$$

The solution for the current in the first loop is

$$i_1 = \frac{v[(Z_1+Z_3+Z_d)(Z_2+Z_4+Z_d) - Z_d^2]}{\Delta} \tag{8-80}$$

The current in the second loop is

$$i_2 = \frac{v[Z_d Z_4 + Z_3(Z_2+Z_4+Z_d)]}{\Delta}$$

Finally, the current in the third loop is

$$i_3 = \frac{v[Z_3 Z_d + Z_4(Z_1+Z_3+Z_d)]}{\Delta}$$

Where in each case, the determinant of the coefficients is

$$\begin{aligned}\Delta = {} & \\ & (Z_3+Z_4)[(Z_1+Z_3+Z_d)(Z_2+Z_4+Z_d) - Z_d^2] \\ & - Z_3[Z_d Z_4 + Z_3(Z_2+Z_4+Z_d)] \\ & - Z_4[Z_3 Z_d + Z_4(Z_1+Z_3+Z_d)]\end{aligned}$$

The equations above constitute the general solution to the problem where each of the impedance elements is a known quantity. In the case of an impedance bridge, the object is to determine an unknown impedance element in terms of three known impedance elements occupying the remaining positions in the bridge, i.e., to determine an unknown, say Z_1, in terms of known values for Z_2, Z_3, and Z_4. In accomplishing this, the bridge must be forced into a balanced condition wherein the current in the detecting element, Z_d, is made equal to zero. If the current in Z_d is to be zero, it is necessary that the current i_2 be equal to the current i_3. As these currents exist in Z_d in the opposite sense, if these currents are equal, the net current in this element will be zero. Upon equating these two currents, the balanced condition is extracted as follows:

$$v\frac{[Z_d Z_4 + Z_3 Z_2 + Z_3 Z_4 + Z_3 Z_d]}{\Delta} = v\frac{[Z_3 Z_d + Z_4 Z_1 + Z_4 Z_3 + Z_4 Z_d]}{\Delta}$$

from which

$$Z_1 = \frac{Z_2 Z_3}{Z_4}$$

One of the keys to constructing a successful impedance bridge is the ability to construct a nearly pure capacitance whose dissipation factor is so small as to be negligible. Such capacitors exist in the form of either carefully constructed silvered mica or polystyrene capacitors. Such capacitors are quite expensive particularly for units that have very exact values and thus are not for universal use.

An examination in detail of a bridge circuit for determining the series resistance and inductance of a physical inductor will conclude this section. Fig. 21 displays a bridge configuration designed to measure the inductance and quality factor of physical inductors.

In the circuit of Fig. 8-21 the resistor R_2 can be selected in decade steps from one ohm to one megohm and constitutes a range selector. Resistors R_3 and R_4 are continuously variable calibrated resistors. The capacitor C, as mentioned previously, is

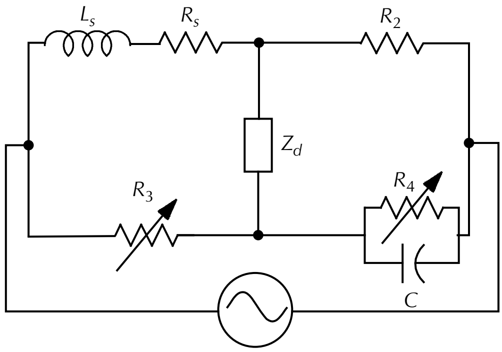


Figure 8-21. An impedance bridge circuit for measuring an unknown physical inductor.

essentially pure with a precise fixed value. The inductor whose properties are to be determined is represented by L_s and R_s. The current indicator is Z_d. The balancing procedure is as follows. Start with the continuously variable resistors set at about the middle of their respective ranges. Step through the possible decade values of R_2 until the value is found which minimizes the indicated current. Next vary R_3 until the current is reduced even further. Then adjust R_4 to obtain an even lower current. At this point, alternately adjust R_3 and R_4 until the current is the least obtainable value. At this point the bridge has reached its null or balanced condition. At this point the following relationships hold.

$$R_s + j\omega L_s = \frac{R_2 R_3}{\left(\dfrac{-R_4 \dfrac{j}{\omega C}}{R_4 - \dfrac{j}{\omega C}}\right)}$$

$$R_s + j\omega L_s = \frac{R_2 R_3}{R_4} + jR_2 R_3 \omega C \tag{8-81}$$

$$L_s = R_2 R_3 C$$

$$R_s = \frac{R_2 R_3}{R_4}$$

$$Q = \omega R_4 C$$

Though differing in details, a similar bridge circuit is possible for measuring the properties of an unknown capacitor. An attentive student should be able to draw such a circuit at this point.

8.7 Constant Resistance Networks

There is an important class of networks called constant resistance networks which are of particular interest to sound system engineers. The members of this class, even though they contain reactive elements, have an impedance which is a constant resistance at all frequencies. Certain members of this class of networks find employment as passive loudspeaker crossover networks. Fig. 8-22 displays the simplest configuration of such a network.

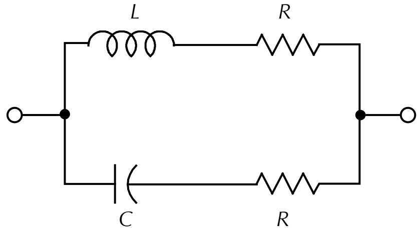


Figure 8-22. Example of a constant resistance network.

The impedance which exists between the terminals of the network of Fig. 8-22 can be made equal to the value assigned to the resistance R provided that L and C bear a particular relationship to each other. This relationship is discovered through the following steps. First, sum the admittances of the upper and lower branches to obtain

$$\frac{1}{R + j\omega L} + \frac{1}{R - \dfrac{j}{\omega C}} = \frac{2R - \dfrac{j}{\omega C} + j\omega L}{R\left(R - \dfrac{j}{\omega C} + j\omega L + \dfrac{L}{RC}\right)} \tag{8-82}$$

Second, inspect the result of the addition and note that if the term in the parentheses of the denominator is adjusted to equal the numerator, then all of the reactive terms cancel out. This adjustment simply requires that

$$\frac{L}{RC} = R \tag{8-83}$$

or

$$\sqrt{\frac{L}{C}} = R \tag{8-84}$$

After this adjustment is made, the total admittance becomes $1/R$ and the impedance then is the reciprocal of the admittance and is simply the value of the resistance.

This particular circuit is often employed as a first order two-way passive crossover network. The constant resistance feature examined above sets one relationship between L and C. The other relationship necessary to allow unique determination of values for L and C comes from the specification of the crossover frequency. In the crossover application, the upper branch forms the low pass filter with the assumption that the low frequency loudspeaker furnishes a load equal to R. (This assumption will be investigated presently.) The lower branch forms the high pass filter with the assumption that the high frequency loudspeaker furnishes an identical resistive load. The crossover angular frequency is that at which the response of each filter is 3 dB down. This requires that

$$\begin{aligned}\omega_0 L &= \frac{1}{\omega_0 C}\\ &= R\end{aligned} \tag{8-85}$$

$$\omega_0 = \frac{1}{\sqrt{LC}} \tag{8-86}$$

As a numerical example, take $R = 8\,\Omega$ and take the crossover frequency to be 800 Hz. The required inductance value is then

$$\begin{aligned}L &= \frac{R}{\omega_0}\\ &= \frac{8\,\Omega}{2\pi 800\ \text{Hz}}\\ &= 1.59\ \text{mH}\end{aligned}$$

Similarly, the required capacitance value is

$$\begin{aligned}C &= \frac{1}{\omega_0 R}\\ &= \frac{1}{2\pi 800\ \text{Hz} \times 8\,\Omega}\\ &= 24.9\ \mu\text{F}\end{aligned}$$

In order for this network's performance to be exact, it is necessary that the load resistance in each branch be a constant 8 Ω. Fortunately, departures from this condition are not serious provided that the departures, if any, occur at frequencies well removed from the desired crossover frequency. A study will be made in the next section dealing with actual loudspeaker behaviors and the corrections required to make loudspeakers serve as loads for passive crossover networks. Additionally, more involved crossover networks will be studied in Chapter 18 *Loudspeakers and Loudspeaker Arrays*.

8.8 Impedance Properties of Moving Coil Loudspeakers

The voice coil of a moving coil loudspeaker certainly possesses resistance associated with the conductor forming the coil and it has self-inductance associated with the coil and the magnetic structure in its environment. When taken together, these two factors constitute the electrical impedance of the loudspeaker when the voice coil is at rest. The resting condition, however, is of absolutely no interest acoustically. The cases of importance are when the loudspeaker is being driven into motion either electrically or mechanically. With electrical drive, the loudspeaker is a sound producer. With mechanical drive, such as from an impinging sound wave, the loudspeaker acts as a microphone. The case of interest here involves electrical drive and sound production. This situation is represented in Fig. 8-23.

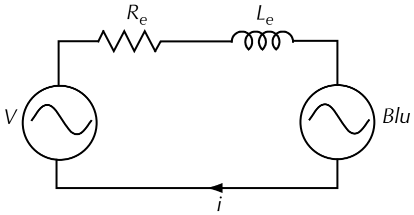


Figure 8-23. Model for determining loudspeaker electrical impedance.

In Fig. 8-23 R_e and L_e are the voice coil resistance and self-inductance, respectively. The voltage source of strength Blu on the right represents the voltage induced in the voice coil as a result of its motion in the magnetic gap of the magnet structure. B is the value of the magnetic induction in the gap measured in teslas. The length of the voice coil conductor in the gap is l measured in meters. The instantaneous velocity of the voice coil is u measured in meters per second. The voltage source on the left is the applied voltage to the loudspeaker and i is the current which results from this applied voltage. The electrical impedance to be determined is that as viewed by the voltage source on the left and is the ratio of v to i. The voltage source on the right is an induced voltage and consequently is subject to Lenz's law. Lenz's law states that an induced voltage acts so as to oppose the cause which produces it. The primary cause in this case is the current i which has a positive sense in the direction of the arrow in Fig. 8-23. Therefore the current in the circuit is given by

$$i = \frac{v - Blu}{R_e + j\omega L_e}$$

Without knowledge of u, this would amount to a dead end. Fortunately, it is possible to make an independent statement about u that also involves the current. The applied force experienced by the voice coil results from the interaction of the current existing in the conductor residing in the magnetic gap with the magnetic field existing in the gap. This force has the form

$$F = Bli \tag{8-87}$$

In mechanical systems, it is possible to draw strong analogies with electrical systems in those instances where the governing equations have the same mathematical structure as the structure exhibited in the electrical case. In the electrical case, the complex electrical impedance is defined to be the ratio of the complex applied voltage to the resulting complex current. In the mechanical case, the complex mechanical impedance is defined to be the ratio of the complex applied force to the resulting complex velocity. Mechanical impedances have the dimensions of kilograms per second. This combination is called a mechanical ohm. The mechanical impedance thus appears as

$$Z_m \equiv \frac{F}{u} \tag{8-88}$$

Assuming, for the moment, that the mechanical impedance is known, it is possible to solve for u in terms of i to obtain

$$u = \frac{Bli}{Z_m} \tag{8-89}$$

This can now be substituted in the first equation above involving i and u to obtain

$$\begin{aligned} i &= \frac{v - \frac{B^2 l^2 i}{Z_m}}{R_e + j\omega L_e} \\ &= \frac{v}{R_e + j\omega L_e + \frac{B^2 l^2}{Z_m}} \end{aligned} \tag{8-90}$$

It is now a simple matter to solve for the electrical impedance of the loudspeaker when in operation. The last equation is solved for the ratio of v to i to obtain

$$\begin{aligned} Z_e &= \frac{v}{i} \\ &= \left(R_e + j\omega L_e + \frac{B^2 l^2}{Z_m}\right) \end{aligned} \tag{8-91}$$

The question at this point becomes "What is the structure of Z_m?" In order to answer this question it is necessary to further explore the electrical-mechanical analogy. Table 8-1 contains one possible set of analogies which is drawn from the mathematical structure of the differential equations describing the behaviors of electrical and mechanical systems.

Table 8-1. Analogies Drawn from the Mathematical Structure of the Differential Equations Describing the Behaviors of Electrical and Mechanical Systems

Electrical	Mechanical
Applied voltage	Applied force
Charge	Displacement
Current	Velocity
Time derivative of current	Acceleration
Resistance	Kinetic friction constant
Inductance	Mass
Capacitance	Compliance or reciprocal stiffness (spring constant)

A moving coil loudspeaker has a mass associated with its motional elements. The suspension elements of the loudspeaker such as the spider and surround act like springs. If the back side of the loudspeaker is enclosed in a closed box, the air in the box acts like a spring having a stiffness inversely proportional to the air volume in the box. The sound energy carried away by sound waves represents a loss of mechanical energy and thus appears as an additional friction constant or mechanical resistance. This particular mechanical resistance is called the radiation resistance. The air mass in direct contact with and in the immediate vicinity of the loudspeaker also adds to the moving mass of the loudspeaker. This effect is represented by the radiation reactance. When all of these factors are taken into account, the mechanical impedance appears as

$$Z_m = R_m + R_r(\omega) + jX_r(\omega) + j\left(\omega M - \left[\frac{K_s + K_b}{\omega}\right]\right) \tag{8-92}$$

Additional explanation of the terms appearing in the expression for the mechanical impedance is useful. The term R_m is the mechanical resistance which accounts for the frictional losses occurring in the suspension when the loudspeaker is in motion.

The term $R_r(\omega)$ is the mechanical resistance associated with radiation energy losses. Unlike R_m, however, this term is frequency dependent with a rather involved frequency dependence, the exact nature of which is not important for the present discussion. The term $X_r(\omega)$ is the radiation reactance which also has an involved frequency dependence. The term in the parentheses describes the resonant interaction between the moving mass and the stiffness of the suspension along with the stiffness of the air in the box. This resonance ordinarily occurs at a low frequency in which case the radiation reactance simply adds a small amount to the term ωM, thus in effect slightly increasing the moving mass. Fig. 8-24 is a plot of the magnitude of the total electrical impedance typical of a 10 inch cone type moving coil loudspeaker.

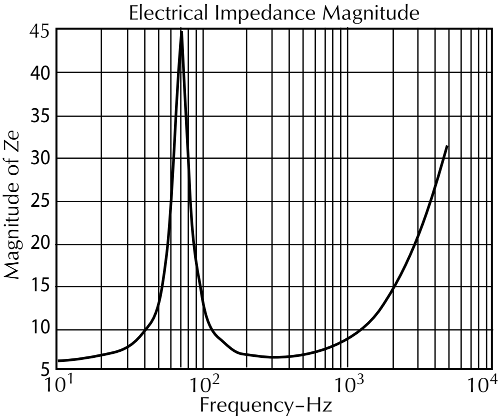


Figure 8-24. Electrical impedance magnitude typical of a moving coil loudspeaker.

The effect of the mechanical resonance of this loudspeaker is clearly displayed in Fig. 8-24 as occurring at a frequency of 70 Hz with an impedance magnitude of nearly 45 Ω. The nominal impedance of this loudspeaker is the minimum value above mechanical resonance and is a little over 6 Ω. This value is not to be confused with the dc resistance of the voice coil which is always lower. If this loudspeaker is to be employed in a two-way system with a passive crossover at 1000 Hz or above, it is necessary to add a network in parallel with the loudspeaker which will render the overall impedance to be resistive and constant in value. Above the minimum in the curve, the impedance magnitude increases with frequency as one would suspect from the voice coil inductance. In this region there is only a minor influence from the mechanical impedance and this influence is diminishing as the frequency increases. The loudspeaker itself behaves as a combination of a series resistance and inductance. The corrective network to be placed in parallel then must be a series resistance and capacitance as was learned from the discussion of constant resistance networks. In arriving at values for this network, one selects a resistance which equals the minimum value on the impedance curve. Call this value R_{nom}. Then one solves for a capacitance using the constant resistance network equations to obtain

$$C = \frac{L_e}{R_{nom}^{\ 2}} \qquad (8\text{-}93)$$

The resistance R_{nom} and the capacitance C are then connected in series with the combination subsequently connected in parallel with the loudspeaker terminals. The impedance curve of the combination is then measured. If the measured impedance is now flat with zero angle through the desired crossover region, the correction is proper and no further work is required. If this is not the case, small modifications are alternately made to both R_{nom} and C until the desired performance is required. Figs. 8-25 and 8-26 display the magnitude and impedance angle behaviors for a properly corrected system.

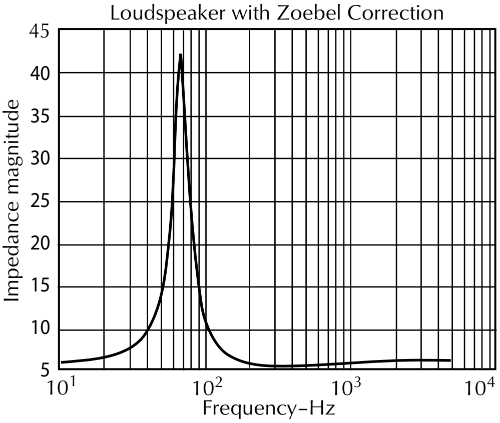


Figure 8-25. Impedance magnitude after correction.

It is important to emphasize that the presence of the correction network in parallel with the loudspeaker terminals does not alter either the electrical or acoustical performance of the loudspeaker in any way. Instead, the combination of the loudspeaker and the parallel connected correction network now constitutes a proper resistive load to properly terminate a passive constant resistance crossover network. The corrective network is called a Zoebel network to honor one of the pioneers of electric filter technology. In Fig. 8-25, it is apparent that the impedance magnitude is constant above 1000 Hz. Equally important, Fig. 8-26 illustrates that the

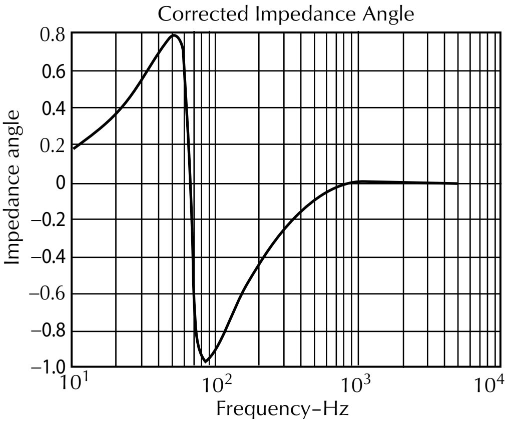


Figure 8-26. Impedance angle after correction.

angle of the overall impedance is zero, that is purely resistive above 1000 Hz.

8.9 Network Theorems

Much reference has been made to voltage and current sources as well as passive circuit elements in all of the previous work. Voltage sources and current sources are also circuit elements, which are further classified as being active circuit elements. The distinction between an active circuit element and a passive one is that the active circuit element contains or controls a source of electrical energy. It is true that capacitors and inductors can temporarily store energy but they do not inherently contain independent sources of energy and hence are classified as passive elements.

Active elements or networks must contain devices or agencies which have the ability, in the thermodynamic sense, to reversibly convert some other form of energy into electrical energy. The amount of energy so converted in the device or agency per unit of charge passing through the device or agency has historically been called the electromotive force or emf. The appearance of the word force in the historical terminology is a misnomer as energy rather than force is the applicable concept. An emf is measured in volts and a volt is a joule of energy per coulomb of electric charge. In the case of an emf, the energy referred to is the amount of some other form of energy converted into electrical energy per unit of electrical charge. Potential difference, previously referred to as simply voltage, is also measured in volts. Potential difference, however, represents the change in electrical potential energy per unit of electrical charge experienced in moving between two points. The letter E will be used here to represent a sinusoidal emf. The letter v will continue to be used to represent a time dependent potential difference.

In order to sustain a current in a closed conducting circuit, the circuit must contain a source of electrical energy provided by an emf. The only exception to this statement occurs in a closed superconducting coil. Even in the superconducting coil case, an emf is necessary in order to first establish a direct current in the coil. Sources of emf appear in many forms. A mechanically driven alternating current generator is a source of emf in which a portion of the mechanical energy of rotation of the generator shaft is converted into electrical energy. Chemical electric cells of both the primary and secondary categories are sources of a dc emf. In this instance an exoergic chemical reaction in the cell makes electrical energy available. A solar panel is a source of a dc emf. Here a portion of the luminous energy from the sun impinging on the cell is converted into electrical energy. One could easily list many other examples. A common property of all of these sources of emf is that they must possess a conducting path between their terminals and along this path will occur losses. For example, the coils in a mechanically driven alternating current generator will possess both resistance and inductance and perhaps significant distributed capacitance. This means that any source of emf will also have some impedance internal to its structure. This situation is depicted in its simplest form in Fig. 8-27.

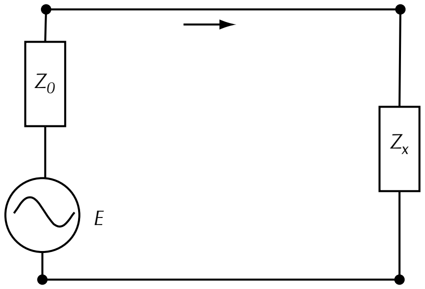


Figure 8-27. A source of alternating emf and its externally applied load.

The elements depicted on the left in Fig. 8-27 represent a physically realizable source of emf. The element on the right in Fig. 8-27 represents the impedance of some circuit which is connected across the terminals of the source of emf. The source of emf is represented by two elements. A circle containing one period of a sine curve, this being the representation of an ideal source of alternating emf (one without losses) and a rectangle representing the internal impedance of the conducting path through the source of emf. The internal impedance of the source of emf is Z_0 and the impedance of the external circuit is Z_x. The arrow indicates the positive sense of the current in the circuit. The sustained current which will exist in a simple conducting loop is the algebraic sum of the emfs (if there are more than one) divided by the total series impedance of the loop including all source internal impedances. In this instance the current is given by

$$i = \frac{E}{Z_0 + Z_x} \tag{8-94}$$

The first circuit theorem to be discussed is the *Maximum Power Transfer Theorem*. This theorem deals with the proposition that given a particular source having a fixed emf and a fixed internal impedance, what value of load impedance must be connected to the terminals of the emf such that the power dissipated in the external load will be a maximum? In answering this question, first rewrite the equation for the current in the circuit of Fig. 8-27 by explicitly displaying the complex nature of both the internal and external impedances.

$$i = \frac{E}{(R_0 + jX_0) + (R_x + jX_x)} \tag{8-95}$$

Earlier it was learned that the power dissipated in a load is directly proportional to the square of the current multiplied by the real part of the load impedance. The next step is to examine the expression for the current and inquire what adjustment can be made to the load impedance that will not change its real part but will make the denominator in the current expression smaller and consequently, the current larger. Upon recalling that reactances can be both positive and negative, the denominator can be made smaller by having the reactance of the load be just the negative of the reactance of the source thus yielding a net reactance of zero. The current now becomes

$$i = \frac{E}{R_0 + R_x} \tag{8-96}$$

The average power dissipated in the external load is then

$$P = 0.5\frac{{E_m}^2}{(R_0 + R_x)^2}R_x \tag{8-97}$$

where,
E_m is the amplitude of the sinusoidal emf.

The next and final step is to find the value of R_x which will make the power expression a maximum. Readers familiar with the differential calculus would differentiate the power expression with respect to R_x and set the derivative equal to zero thus obtaining a simple algebraic expression which can easily be solved for the magic value of R_x satisfying the problem. Those readers not familiar with calculus can still arrive at the answer after the expenditure of some additional effort. It is only necessary to let R_x be a variable expressed in fractions and multiples of R_0 and graph the function

$$G = \frac{R_x}{(R_0 + R_x)^2} \tag{8-98}$$

A carefully drawn graph will have a maximum occurring when R_x equals R_0.

The conclusion is that the power in the external load is a maximum when the load impedance is the complex conjugate of the source impedance. As a reminder, complex conjugate means equal real parts with imaginary parts equal but opposite in sign.

It is important to note, however, that when the power dissipated in the load has been maximized, the power transfer efficiency is only 50% as an equal amount of power is being dissipated within the source itself. Additionally, under these conditions, the voltage applied to the load is only one-half of the open circuit voltage of the source. This type of match is often desirable in communications systems where a premium is placed on signal power. It is not employed in commercial power systems where a premium is placed on voltage regulation and overall efficiency. Note also, this theorem can not be applied to audio power amplifiers because the emf associated with the power amplifier output is not a constant as required by the theorem, but rather is a function of the load impedance to which it is connected. Power amplifiers are purposely designed to have low output impedances (high damping factors) and are designed to work into load impedances equal to or greater than a specified minimum which is significantly larger than the output impedance.

The second theorem to be discussed is *Thévenin's Theorem*. This theorem deals with the proposition that a linear network of any number of emfs and any combination of impedances which communicates with the external world through only two terminals is equivalent to a single emf E_0 in series with a single impedance Z_0. E_0 is equal to the voltage present between the actual network's terminals when it is disconnected from the external world and Z_0 is the impedance measured between the actual network's terminals when it is disconnected from the external world and all internal emfs are inactive. The value of this theorem and an elaboration of the terminology used therein can best be explored through an example problem.

Fig. 8-28A displays a circuit problem where the objective is to determine the current which exists in the physical inductor located in the central branch. The circuit contains two sinusoidal generators each having a time dependence of $\cos(\omega t)$ and each having internal resistance as indicated. The emfs associated with the generators are the rms values. The inductor is connected between the points labeled a and b. One could solve this problem by assuming mesh currents similar to the technique employed in the section dealing with the impedance bridge. One would then have to write two simultaneous equations, solve them, and then use knowledge of the mesh currents to determine the net current in the central branch. The Thévenin approach is simpler in this instance. First, disconnect the inductor between a and b and find the potential difference or open circuit voltage which exists between a and b under these conditions. This circuit appears in Fig. 8-28B. In Fig. 8-28B the net emf is 10 V − 5 V = 5 V and the total series impedance is 2 Ω + 3 Ω = 5 Ω. The current in the loop is then 5 V divided by 5 Ω or 1 A. From the generator on the left, V_{ab} is then 10 V less the drop across its internal resistance which is 1 A times 2 Ω giving 10 V − 2 V = 8 V. Now one determines the impedance which exists between a and b when the emfs of the generators are set equal to zero but with the generators still in place. In this instance, one has 2 Ω paralleled by 3 Ω giving a value of 1.2 Ω. The Thévenin equivalent of the outer loop is then a generator having an emf of 8 V and an internal impedance which is a pure resistance of 1.2 Ω. Finally, in Fig. 8-28C, this source is connected to the load inductor and it is a simple matter to determine the current. The impedance of this series loop is

$$
\begin{aligned}
(1.2 + 1.8 + j4)\,\Omega &= \sqrt{3^2 + 4^2}\,\Omega e^{j\,\mathrm{atan}\frac{4}{3}} \\
&= 5\,\Omega^{j0.927}
\end{aligned}
$$

The current in this loop which is the current in the inductor has an rms value of 8 V divided by 5 Ω or 1.6 A and the current lags behind the voltage of the generator by an angle of 0.927 radian.

In summary, Thévenin's theorem says that any linear two terminal active network can be replaced by an ideal source of emf in series with an impedance. The value of the emf is equal to the open circuit voltage of the network between the two terminals and the impedance is the measured impedance between the terminals when the sources internal to the network are inactive.

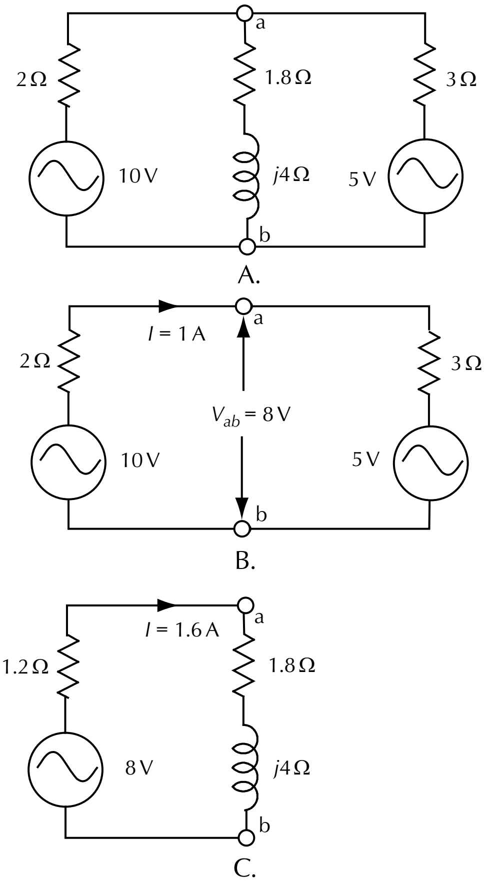


Figure 8-28. Application of Thévenin's theorem.

A related network theorem is called *Norton's Theorem*. Norton's Theorem states that any linear two terminal active network can be replaced by an ideal current source paralleled by an admittance. The value of the current source is the short circuit current which could exist between the two terminals of the actual network and the admittance is the admittance measured between the two terminals of the network when the sources internal to the network are inactive. The Norton equivalent which could replace the Thévenin equivalent of the sample problem can be derived from the Thévenin equivalent and is shown in Fig. 8-29. The Norton values are determined as follows. When a is shorted to b in the Thévenin equivalent, the current is 8 V/1.2 Ω = 6⅔ A. The admittance between a and b with the voltage set equal to zero in the Thévenin equivalent is the reciprocal of 1.2 Ω or ⅚ siemen (S).

In order for each and every one of the foregoing network theorems to be valid, it is necessary that the circuit or system be linear. Being linear means that the laws of physics governing the system appear in the form of linear differential equations. One of the properties of such equations is that if there exist

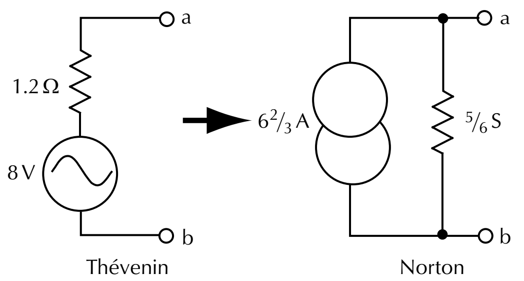


Figure 8-29. Norton equivalent derived from the Thévenin equivalent.

several solutions to a particular linear differential equation, then a sum of these solutions is also a solution to the equation. This leads to what is termed the principle of superposition or the *Superposition Theorem*. In a network containing several sources, either voltage or current, and several impedances, one applies superposition by calculating the effects of each source acting individually with all other sources set equal to zero. This is done for each source in turn. When this is done, it is important to note that even though a source may be inactive, its impedance still remains in the circuit. The final solution when all sources are active simultaneously is just the sum of the solutions obtained when the sources are acting one at a time. As an example, superposition will be applied to the circuit of Fig. 8-28B in order to determine the potential difference between *a* and *b*. Consider that the 10 V source is active with the 5 V source set equal to zero. In this instance, the 3 Ω and 2 Ω resistors form a voltage divider across the 10 V source and V_{ab} is then

$$10\,\text{V} \times \frac{3\,\Omega}{(2+3)\,\Omega} = 6\,\text{V}$$

Now consider that the 5 V source is active with the 10 V source set equal to zero. The situation here has the 2 Ω and the 3 Ω resistors forming a voltage divider across the 5 V source yielding a value of V_{ab} of

$$5\,\text{V} \times \frac{2\,\Omega}{(2+3)\,\Omega} = 2\,\text{V}$$

When both sources act simultaneously, V_{ab} is the sum of these two individual solutions yielding the value of 6 V + 2 V or 8 V. There are many more network theorems, many of which are more specialized. The ones studied here are the ones most often invoked in routine circuit analysis.

8.10 The Technician's Viewpoint

The following section of *Interfacing Electrical and Acoustic Systems* is written from the viewpoint of the technician with a handheld impedance meter seeking to confirm values and adjust them where necessary. We treat the values as resistances which in the case of mixer output, amplifier inputs, etc., are reasonable. Microphones and loudspeakers present more complex values (covered by Dr. Patronis in the first section of this chapter).

8.11 Impedance Defined

What is impedance? Many technicians have put an ohmmeter across the voice coil of a 16 Ω loudspeaker and been surprised to read 4.5 Ω or less. What have they read with the ohmmeter? The dc resistance of the voice coil. Is this the impedance?

No.

Now, let us use a bridge circuit to read the ac resistance (*R*) of the loudspeaker. Is this the impedance? No, but it is part of the impedance.

What is impedance? It is defined as the total opposition, including resistance and reactance, a circuit offers to the passage of alternating current. We all know what resistance is, but what do "opposition" and "reactance" mean? Opposition is a resistance, a restraint, or a hindrance. From this we can conclude that in ac circuits there exists some other resistance-like component, and this additional restraint, which adds to that of the ac resistance, is called reactance.

There are two kinds of reactance, capacitive reactance (X_C) and inductive reactance (X_L). Reactance varies with frequency, whereas ac resistance tends to stay the same with frequency (certainly over the audio range).

Now, let's add two more terms to our collection and proceed to measure impedance. Let us designate $|Z|$ as the magnitude of the impedance and the power factor (*PF*), giving us the following terms which are terms for use in steady state circuits:

1. *R* is ac resistance.
2. *X* is reactance.
3. X_C is capacitive reactance.
4. X_L is inductive reactance.
5. $|Z|$ is the magnitude of the impedance (*Z* is complex).
6. *PF* is power factor or $\cos\theta$ where θ is the angle of the phase difference between the voltage and the current.

If we now measure voltage and current in a real circuit, the current lags behind the voltage in phase when an inductor-like device is in the circuit, and the current leads the voltage in phase when a capacitor-like device is in the circuit. ELI THE ICE MAN helps us remember this relationship: E (voltage) is ahead of I (current) when L (inductance) is predominant and I is ahead of E when C (capacitance) is predominant in the circuit. These are root mean square (rms), voltage and current values.

8.11.1 Making Reactance Visible

By setting up a standard way of plotting resistance and reactances, their action and interaction become obvious. We can do this with rectangular coordinates. Components 90° apart in their phase are represented by vectors that are 90° apart, see Fig. 8-30. From Fig. 8-30, we can write the following equations for impedance:

$$|Z| = \sqrt{R^2 + X^2} \tag{8-99}$$

where,
$|Z|$ is the magnitude of the impedance,
R is the resistance,
X is the total reactance,
$X = (X_L \text{ and } X_C)$.

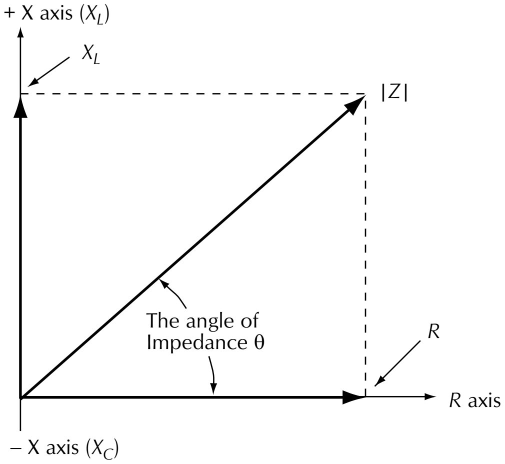


Figure 8-30. Calculation of impedance.

This describes how to plot the action in Fig. 8-30. The impedance angle can be found by:

$$\tan^{-1}\left(\frac{X}{R}\right) = \theta \tag{8-100}$$

where,
θ is the angle of the impedance in degrees or radians.

8.11.2 Impedance Notation

The following equations are also used when working with impedance.

$$\sqrt{R^2 + X^2} = \text{Magnitude of Impedance} \tag{8-101}$$

$$|Z|(\cos\theta) = R \tag{8-102}$$

$$|Z|(\sin\theta) = X \tag{8-103}$$

$$\tan^{-1}\left(\frac{X}{R}\right) = \theta \tag{8-104}$$

$$\frac{R}{\cos\theta} = |Z| \tag{8-105}$$

$$\frac{X}{\sin\theta} = |Z| \tag{8-106}$$

$$\begin{aligned}\frac{R}{|Z|} &= PF \\ &= \cos\theta\end{aligned} \tag{8-107}$$

$$\frac{X}{|Z|} = \sin\theta \tag{8-108}$$

$$\cos^{-1}\left(\frac{R}{|Z|}\right) = \theta \tag{8-109}$$

$$\sin^{-1}\left(\frac{X}{|Z|}\right) = \theta \tag{8-110}$$

where,
$|Z|$ is the magnitude of impedance,
θ is the angle of impedance,
$\cos\theta$ is the power factor,
R is the resistance,
X is the reactance,
Z is the magnitude and angle $= |Z|e^{j\theta}$.

These equations describe how to obtain the magnitude of the impedance vector and the angle between it and 0° on the ac-resistance axis.

In Fig. 8-30 we can see that, when angles are small, the impedance value approaches the ac-resistance value. Conversely, when angles are large, the reactive component must be carefully measured. The performance of transformers, loudspeaker voice coils, and the like exemplifies conditions in which there are large angles. (A loudspeaker has a blocked impedance—cone cannot move—and a motional impedance. When the loudspeaker impedance does not rise or fall with a change in frequency, it is essentially resistive within that frequency range. The ac resistance plus the motional impedance equal the total impedance measured.)

8.11.3 Amplifier and Loudspeaker Impedances

Amplifiers can present outputs for loudspeakers rated for 4 Ω, 8 Ω, 16 Ω, etc. The actual source impedance of modern amplifiers is quite low, usually on the order of 0.1 Ω or less.

Amplifiers also can provide constant voltage output ratings, 25 V, 70.7 V, 100 V, and 200 V, as well as, on occasion, constant current output.

Once the type of interface has been ascertained and the appropriate amplifier chosen, then the next device back towards the input of the system needs to be ascertained. This may be a series of pads, resistive matching networks, or an active device. It is important that:

1. It can output its full output for best signal-to-noise ratio, *SNR*.
2. It is an appropriate match in circuit type, balance, unbalanced, etc., and in circuit impedance.

It has been found over the years that most signal processing equipment should be in the system near the input of the power amplifier. The signal through a sound system should emphasize *SNR* near its input, and spectrum shaping and other frequency-dependent level changes near the output where the signal levels are more robust. We will treat all devices between the mixer and the power amplifier as “black boxes” in order to define the treatment of their inputs and outputs.

8.11.4 Complex Impedance

For those desiring a look at loudspeaker values of real and imaginary parts, angle of impedance, etc., use of a modern analyzer is helpful, in this case a Goldline TEF, see Figs. 8-31 through 8-37.

The published EIA data of reputable manufacturers is used for acoustic devices such as microphones and loudspeakers with predictable results in terms of matching and levels. A system designer not using standard components but rather starting from zero to a system of unique components has a wide horizon of options.

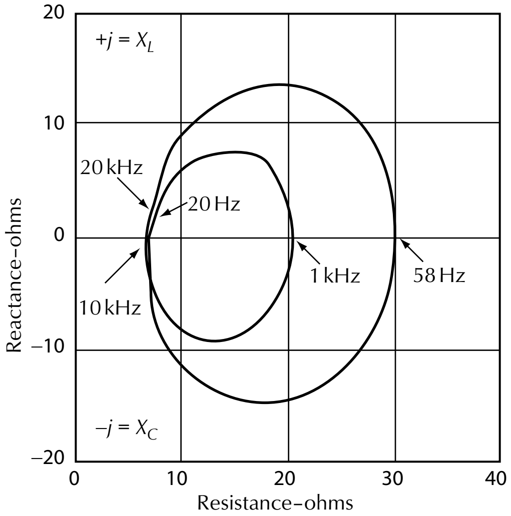
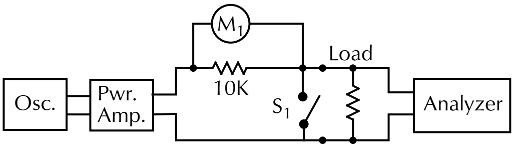


Figure 8-31. Complex impedance Nyquist plot.



1. Close S_1 and set M to 0.1 V.
2. Open S_1 attach test load and measure.
3. Adjust analyzer input sensitivity until you read test load value.
4. Measure desired load.

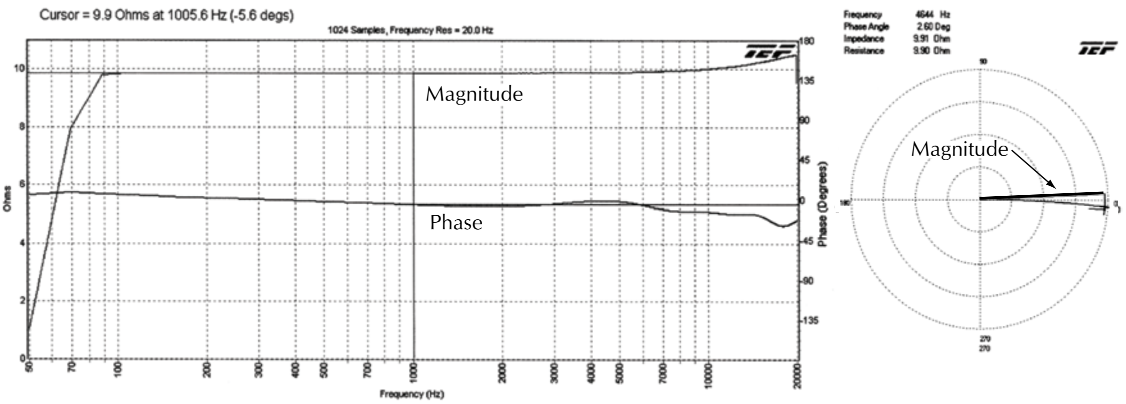
Figure 8-32. How the plots in Fig. 8-33through 8-37 were made.

8.12 Handling the Acoustic Input and Output of the System

Knowing the expected performer’s maximum L_P allows for an intelligent choice of microphone sensitivities. The choice of microphones can be affected by:

1. Appearance.
2. Sensitivity.
3. Reliability.
4. Circuit types—wired, wireless, dynamic condenser, etc.
5. Directivity characteristics.
6. Freedom from wind noise, extraneous electrical field pickup, and handling noise.

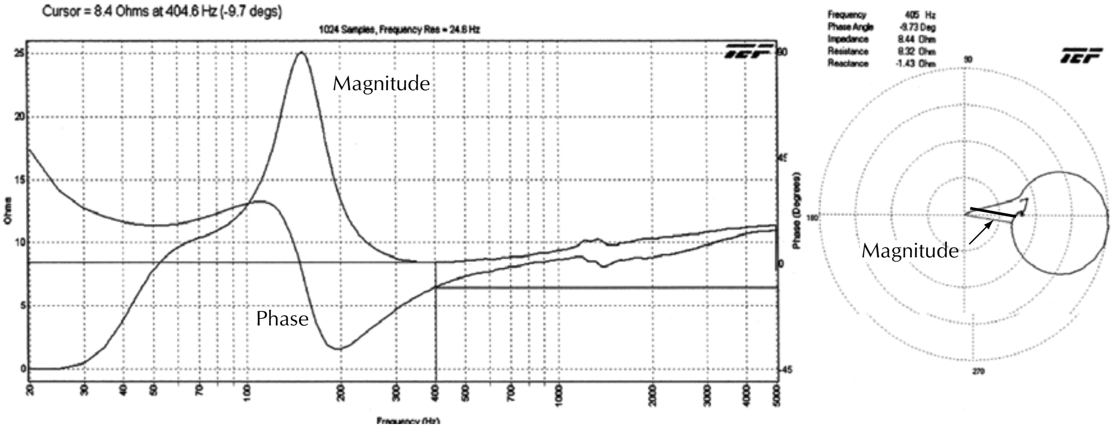
For our purpose of interface, sensitivity and impedance need explanation.



Magnitude and phase response

Nyquist plot

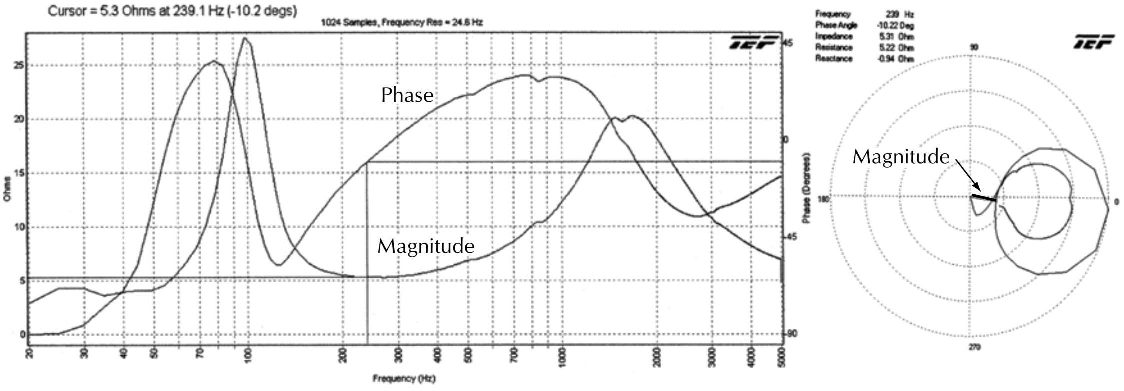
Figure 8-33. A resistive input (10 Ω, 5% resistor).



Magnitude and phase response

Nyquist plot

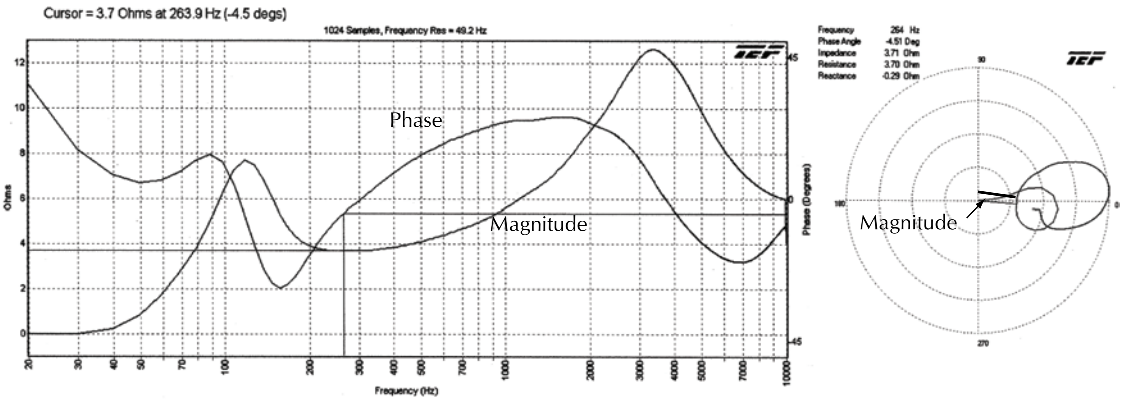
Figure 8-34. A single cone in an enclosure box. Cursor at the lowest value magnitude 8.4 Ω. Manufacturer rating is 8 Ω.



Magnitude and phase response

Nyquist plot

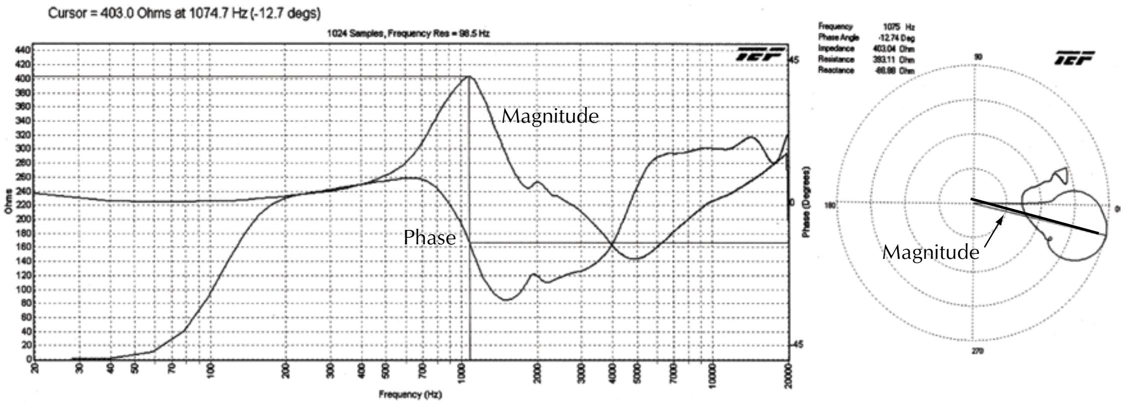
Figure 8-35. A two way system. Cursor at the lowest value 5.3 Ω. Rating chosen by manufacturer is 8 Ω.



Magnitude and phase response

Nyquist plot

Figure 8-36. A smaller two way system. Cursor at the lowest point 3.7 Ω. Rating chosen by manufacturer is 4 Ω.



Magnitude and phase response

Nyquist plot

Figure 8-37. A "rock" performer's microphone specified to operate into an "open" circuit.

8.12.1 The EIA Microphone Rating

The EIA has chosen to rate microphone sensitivity as:

$$G_M = S_V - 10\log R_{MR} - 50\ \text{dB}* \qquad (8\text{-}111)$$

where,

$S_V = 20\log E_O - Test\ L_P + 74$,

R_{MR} is the center value of the impedance range, typically 38 Ω, 150 Ω, 600 Ω, Table 8-2,

E_O is the open circuit voltage at $Test\ L_P$,

L_P is the $Test\ L_P$ (usually either 94 dB or 74 dB),

*50 dB = 10log (1/0.001 + (94 − 74).

To obtain G_M from other ratings use:

$$S_V = 20\log E_O - Test\ L_p + 74 \qquad (8\text{-}112)$$

then insert S_V into the G_M equation (typical $Test\ L_P$ is either 94 dB or 74 dB).

Table 8-2. R_{MR} defined

Ranges (Ω)	**Values Used** (Ω)
20–80	38
80–300	150
300–1250	600
1250–4500	2400
4500–20,000	9600
20,000–70,000	40,000

1 dyn/cm² = 1 µbar = 0.1 Pa = 0.1 N/m²

G_M provides the microphone's output level in dBm (theoretical) for an input L_P = 0 dB. This allows the output of the microphone to be added directly to the performer's L_{PMAX} for the total electrical output of the microphone at the mixer's input.

Low sensitivity microphones, suitable for "rock" concerts, can have an output level, when used by quiet talkers, too near the noise floor—for example, in conference rooms, by ministers, anyone speaking softly. Conversely a very high sensitivity microphone choice for the high level performer can result in input overload at the mixer's input amplifier.

This is an "available input power" figure. An otherwise noiseless device has a thermal noise floor of −132 dBm (for a spectrum from 20 Hz–20,000 Hz). The G_M figure allows an instant estimate of *SNR* at the very beginning of the system; however, the acoustic *SNR* at the microphone is a separate case.

8.12.2 The Mixer Output

The audio engineer needs to know the source voltage and source impedance of the device. This is depicted in Fig. 8-38.

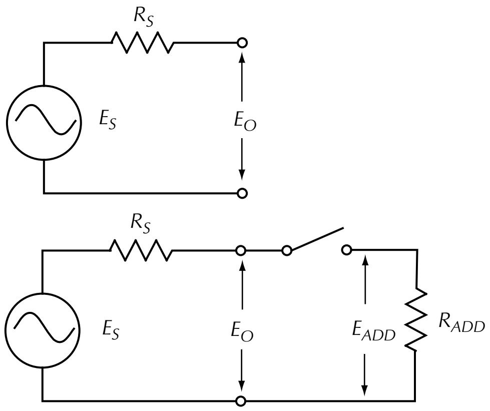


Figure 8-38. Finding R_s from voltage measurements.

Using Fig. 8-38, we can find R_s with the following equation:

$$R_s = R_{ADD}\left(\frac{E_O}{E_{ADD}} - 1\right) \qquad (8\text{-}113)$$

Measure E_O, close switch, measure E_{ADD}

Example

Let,

$R_{ADD} = 600\,\Omega$,

$E_O = 1.0\,\text{V}$,

$E_{ADD} = 0.8\,\text{V}$,

$$R_S = 600\left(\frac{1.0\ \text{V}}{0.8\ \text{V}} - 1\right)$$
$$= 150\,\Omega$$

The source voltage can be found with the following equation:

$$E_S = E_{IN}\left(\frac{R_S + R_{IN}}{R_{IN}}\right) \qquad (8\text{-}114)$$

8.12.3 Available Input Power

From this we can find the available input power in dBm to the input of the device following the mixer

$$AIP \text{ in dBm} = 10\log\left(\left[\frac{(E_S)^2}{0.001 R_S}\right] \times \cos\theta\right) - 6.01\text{ dB} \qquad (8\text{-}115)$$

Each subsequent device is handled in exactly the same manner. We only care about *AIP* from the previous device and the *AIP* at the output of the device itself, in order to specify the gain or loss the device occasions, Fig. 8-39.

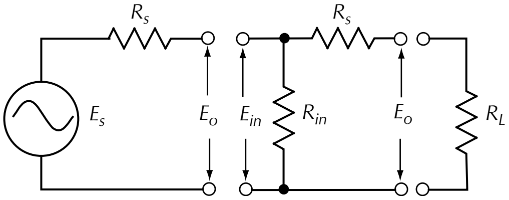


Figure 8-39. AIP of a system.

At the output of the power amplifier we compute the actual power into the load rather than its *AIP*

$$P = \left(\frac{(E_{OUT})^2}{Z_{LOAD}}\right)(\cos\theta) \qquad (8\text{-}116)$$

or

$$10\log\left(\frac{P}{0.001\text{ W}}\right) = dBm \qquad (8\text{-}117)$$

where,

P is the power in watts.

8.12.4 Open and Matched Circuits

Fig. 8-40 shows an open circuit and a matched circuit. For each circuit:

$$E_L = E_s\left(\frac{R_L}{R_L + R_s}\right) \qquad (8\text{-}118)$$

where,
E_L is the load voltage,
E_s is the source voltage,
R_L is the load resistance.

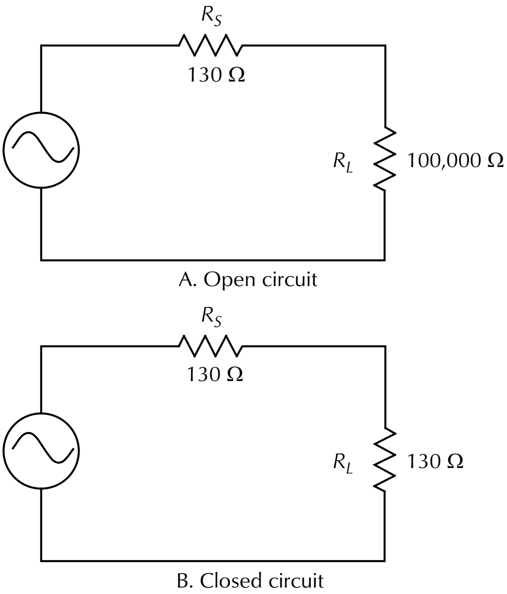


Figure 8-40. Comparison of open and matched circuit.

In the circuits of Fig. 8-40, assume the voltage E is 1 V in both cases. The open circuit load voltage, Fig. 8-40A would be:

$$E_L = 1.0\left(\frac{100{,}000}{100{,}130}\right)$$

$$E_L \cong 1\text{ V}$$

For the matched circuit, Fig. 8-40B:

$$E_L = 1.0\left(\frac{130}{260}\right)$$

$$= 0.5\text{ V}$$

The ratio of the two load voltages, converted to decibels, is:

$$20\log\frac{0.5}{1} = -6\text{ dB}$$

8.12.5 When to Measure Z

Knowing the actual source, input, output, and load |Z| is necessary if systems are to be installed properly. The installer must be aware of the difference between the values normally specified by the manufacturers of equipment and the actual measured values needed by the installer for matching.

There is usually a rated R_s, R_{IN}, R_{OUT}, and R_L as well as an actual value. R_s is the actual source impedance (often an electroacoustic transducer) and R_{IN} is the actual input impedance of a system device. R_{OUT} is the actual output impedance of a system device (as distinguished from its output impedance rating) and R_L is the load impedance (output impedance ratings are usually the desired R_L). Rated input impedances are often the desired R_s.

The term "matching" may be read as "appropriate match." Normally, only in the case of passive devices is the appropriate value also the exact value.

There are two types of reactance—inductive and capacitive. Resistance also has two components—ac and dc. Impedance also can be seen as being composed of lumped parameters (circuit components all in one place) or distributed parameters such as 100 miles of telephone cable.

8.12.6 System Problems Located by Z Measurements

Problems detectable via *Z* measurements vary. They can detect such problems as reactive 70 V transformers overloading power amplifiers, woofers in incorrectly ported enclosures, link circuits not matching passive devices to be inserted in them, and discovery of intermittent circuits.

8.12.7 Gain and Loss Blocks*

Gain blocks are available in increments as small as 30 dB and as large as 100 dB. Very useful gain increments fall in the 40 dB area. If a high *SNR* is to be maintained with variable gain blocks (e.g., mixers and power amplifiers with gain control) while a proper peaking factor and low distortion are preserved, the variable gain control must be properly set. In the case of fixed gain blocks, variable or fixed loss blocks will need to be inserted as required.

Loss blocks can be attenuators, mixing networks, equalizers, and pads. A rule of thumb for a typical sound system is that after algebraically totaling all the gains and losses, you should have an overall gain figure of approximately 115 dB. For example, assume a microphone is calculated to have a sensitivity level of −59 dBm in a sound field with an L_P of 94 dB receives an acoustic input of 88 dB. Then 94 dB – 88 dB = 6 dB, and –59 dBm – 6 dBm = –65 dBm. If the sound system has a 100 W output (+50 dBm), then we need 115 dB of gain to get from the level at the mixer input to the level at the loudspeaker at full power.

---

*Gain describes what "level" change occurs at the system output upon insertion of the device into the system. If the level goes down it is a "loss". If the level goes up it is a "gain".

8.12.8 Typical Mixer Amplifier

A typical mixer amplifier has the following specifications:

Gain	87 dB
Power output	+18 dBm (with low distortion)
Output noise	80 dB below full output

We know we must allow 10 dB as a meter lag factor, so the –65 dBm program level out of the microphone should cause the output of the mixer to reach +8 dBm. Therefore, we need to adjust the overall gain of the mixer to 73 dB. (The attenuators in the mixer can be set back approximately 87 dB – 73 dB = 14 dB of working loss.) Suppose, for the moment, that we are going to connect the output of this mixer directly to the input of a power amplifier having the following characteristics:

$$\begin{aligned} \mathit{Power\ output} &= 100\ \mathrm{W} \\ &= 50\ \mathrm{dBm} \\ \mathrm{Gain} &= 64\ \mathrm{dB} \end{aligned}$$

8.13 Total Electrical Gain of a System

The total electrical gain of a system is found using the following equations:

$$\begin{aligned} \mathit{Gain\ in\ dB} &= 20\log\frac{E_{OUT}}{E_O} + 20\log\left(\frac{R_{IN}}{R_S + R_{IN}}\right) \\ &\quad + 10\log\left(\frac{R_S}{R_L}\right) + 6.02\ \mathrm{dB} \end{aligned} \tag{8-119}$$

where,

Gain in dB is the electrical gain of the total system,

$20\log\left(\frac{E_{OUT}}{E_O}\right)$ is the voltage amplification,

$20\log\left(\frac{R_{IN}}{R_S + R_{IN}}\right)$ is the coupling factor,

$10\log\left(\frac{R_S}{R_L}\right)$ is the impedance mismatch,

6.02 dB is the difference between a matched circuit and an open circuit.

This means the amplifier will reach full output from 50 dBm – 64 dB = –14 dBm. Again, to provide our 10 dB meter lag factor, the most output we would want to see at the output of the power amplifier would be +40 dBm (10 W) of program material. Therefore, the maximum input power would be:

$$-14\ \mathrm{dBm} - 10\ \mathrm{dB} = -24\ \mathrm{dBm}$$

The mixer amplifier is putting out a program level of +8 dBm, and it should continue to do so to ensure the maximum *SNR*. Consequently, we need to insert +8 to –24 dB (32 dB) of attenuation in the form of a pad or an input attenuator. This set of circumstances is illustrated in Fig. 8-41. Note that the pad has replaced the gain overlap.

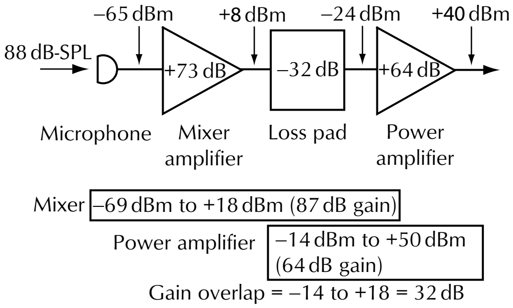


Figure 8-41. Simple system for gain overlap.

8.13.1 A More Complicated System

Fig. 8-42 illustrates a more complex system, and the graph below the block diagram is an example of a gain chart. When first inspecting the total losses in the system (10 dB + 14 dB + 14 dB + 14 dB = 52 dB), you should recognize that you had exceeded the gain overlap available in the mixer and amplifier of the previous example. This immediately suggests that you need a line amplifier. It is highly desirable never to come closer than 10 dB to the original input level at any point in the system beyond the input. This practice assures audio designers that the noise voltages do not become additive—two equal noise levels add to increase the noise 3 dB. For this reason, the required gain block is placed between two loss blocks rather than following the last gain block. Observe, too, that the second gain block is also adjusted to its highest program level, which leaves it a 10 dB meter lag factor.

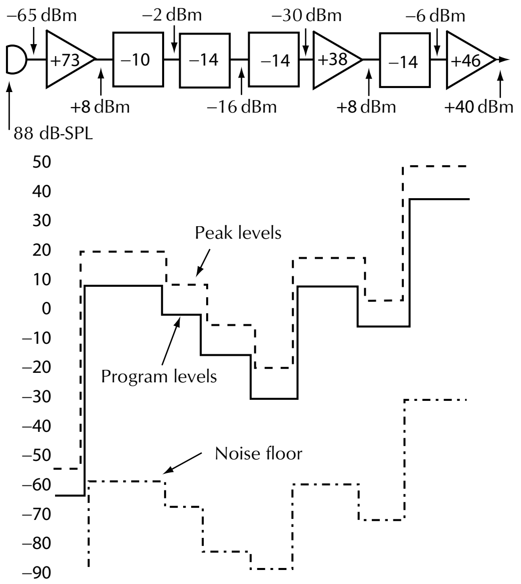


Figure 8-42. More complex system.

8.14 Interfacing the Electrical Output Power to the Acoustic Environment

The final electrical signal power in dBm can be directly added to the loudspeaker's EIA sensitivity rating in dBm (i.e., the 30 ft, 0.001 W rating). The EIA standards, both for microphones and for loudspeakers, wherein L_Ps are directly converted to dBm and the reverse, offer the quickest, easiest, and surest technique for systems engineers. While component-oriented individuals occasionally express a desire for a voltage system rather than a power system, they almost totally ignore the requirement of converting from the acoustic to electrical domain and back again to the acoustic. The very real final product of a sound system is acoustic power and its equitable distribution to listeners and a significant diminution of it to nonlistener areas. Certainly a systems approach can be constructed along lines other than power, but we will be surprised if a more logical and consistent technique than that of the dBm method is arrived at by such advocates.

8.14.1 Simplified Efficiency Calculations

A 100% efficient radiator for a Q equal 1.0 at a radius of 0.282 m and an electrical input of 1.0 W would produce an acoustic level of $Lp = 120$. Therefore the same radiator would produce

$$120\text{ dB} + 20\log_{10}\frac{0.282\text{ m}}{1.0\text{ m}} = 109\text{ dB}$$

for 100% efficiency from 1.0 W at 1.0 m, with a Q = 1.0.

A radiator with100% efficiency from 0.001 W at 30 ft with a Q equal 1.0 would produce

$$120\text{ dB} + 20\log_{10}\frac{0.282\text{ m}}{9.144\text{ m}} + 10\log_{10}\frac{0.001\text{ W}}{1.0\text{ W}}$$

$$= 59.78\text{ dB}$$

The following are the changes in SPL with varying Q at two different references.

1.0 W, 1.0 m	Q	0.001 W, 30 ft	Q
109 dB	1.0	59.78 dB	1.0
112 dB	2.0	62.78 dB	2.0
115 dB	4.0	65.78 dB	4.0
118 dB	8.0	68.78 dB	8.0
121 dB	16.0	71.78 dB	16.0

3.281 ft = 1.0 m.
To change from 0.001 W to 1.0 W add 30 dB.
To change from 1.0 m to 30 ft subtract 19.22 dB.

8.14.2 Damping factor[**]

A loudspeaker is intended to transfer energy through air, by setting the air in motion alternately creating high-and low-pressure zones near the cone. In order to move air, the speaker cone must move. The cone has some non-zero mass, so it has kinetic energy which is non-zero. The energy is stored when the cone is accelerated, and given up when it slows down. Because of this storage, the cone does not instantly reach full speed when a "step" of voltage is applied to the voice coil, nor does it instantly stop and reverse direction when the polarity of the voltages is reversed. When the cone is moving, its kinetic energy must be removed in order to stop it. That is, since work equals force times distance, and also work done on the moving body equals the change in kinetic energy of the body, a force must be applied while the cone moves some distance, in order to stop it. The force applied may be due to suspension friction, it may be due to the air resisting cone motion, or it may be due to electrical damping. If the rate of energy removal is low, (low damping, low friction, and poor coupling to the air,) oscillations may occur at the resonant frequency of the speaker – the characteristic "hangover" or "tubbiness" in the bass, (ringing also occurs in tweeters, and is often controlled by adjusting the suspension friction.) The damping force "F" is dependent on the current, "I" which is inversely proportional to the sum of the resistances, including the voice coil resistance! The voice coil resistance is the largest resistance in the circuit typically 5–7 Ω for nominal 8 Ω loudspeaker, it is the determining element.

[**] Damping Factor article was published in the *Syn-Aud-Con Newsletter*, Vol.15, No. 2, P 30. Winter 1988 with permission from Gerald Tiers. The article here is an edited version of the Newsletter article.

Often neglected is an additional resistor which is the acoustic radiation resistance. This provides additional damping, varying from little, in an inefficient acoustic suspension speaker, to a possibly dominant effect in a high-efficiency horn system. The horn system is likely to be very tolerant of low damping ratios, as the cone is coupled to a greater mass of air, so that its own mass becomes less important by comparison. The effect is similar to suppressing the "ringing" of an L-C filter by properly terminating the output. Speakers generally have an optimum amount of damping, with which they perform best, and the designer picks a trade-off involving efficiency, size, range, and then adjusts the damping to conform to the desired filter function response. To conclude:

Electromagnetic forces in the loudspeaker are a function of the current in the voice coil. For a given voltage, the current in the coil is determined by the total circuit resistance, including that of the coil. It is incorrect to leave voice coil resistance out. Those wonderful damping ratios in the 30s and 40s are wholly imaginary, and those amplifier damping ratios of 1,000 will not be attainable in practice.

High-efficiency horn loaded loudspeakers are necessary in large venues. The least number of loudspeakers introduced into a system that adequately covers the audience is good engineering practice. The famous Klipschorn utilized the corner of a room in order to horn load its low-frequency driver. The Klipschorn's vastly superior transient response to low frequencies was demonstrated to me when I connected one to a Hammond electronic organ followed by hearing the clicks every time a key was depressed. The performance of the Hammond organs of that era was entirely dependent upon loudspeakers with much uncontrolled mechanical damping. The Klipschorn was carefully designed for reproducing accurate sound whereas the speaker in the Hammond was designed to produce sound.

Our thanks for much of the above to Jerrold S. Tiers, Clayton, Mo. the key facts appeared as a "Letter to the Editor" of a popular audio magazine, correcting an erroneous article that had appeared there.

8.14.3 Electrical Gain of a Typical System

When insertion gains and losses of sound system components have been correctly measured, their individual values give the level changes we will read on a sound level meter at the listener's ears in the auditorium if we remove the component in question.

8.14.4 Treating Equalizer Loss

The passive equalizer poses an interesting problem in that its loss is frequency-dependent. We suggest taking an estimated working loss value of 10–14 dB if the sound system is to be a reinforcement system and 4–6 dB if it is a playback system. In this example, we would take a 10 dB loss. (Active devices include gain makeup internally.)

8.14.5 Example of Gains and Losses

The following is an example of how to find the gains and losses and outputs of various devices in a sound system.

Microphone

Given:

$E_o = 0.008\,\text{V}$ (open circuit),

$L_P = 94\,\text{dB}$ (test *SPL*),

$R_{mr} = 150\,\Omega$ (manufacturers specified impedance),

$L_m = 70\,\text{dB}$ (talker's level at microphone).

Find:

$$\begin{aligned} S_V &= 20\log E_o - L_P + 74 \\ &= -61.9\text{ dB re }1.0\text{ V} \end{aligned}$$

$$\begin{aligned} G_m &= S_V - 10\log R_{mr} - 50 \\ &= -133.7\text{ dBm for } L_P = 0 \end{aligned}$$

$$G_m + L_m = -63.7\text{ dBm at the mixer input}$$

Mixer

Given:

$dBm\ in = -63.7$ dBm (output of preceeding device),
$E_S = 9.1$ V (output of this device),
$Z_S = 150$ (output impedance of this device),
$\theta = 0°$ (angle of impedance).

Find:

$$L_{AIP} = 10\log\left[\frac{E_S{}^2}{0.001Z_S} \times \cos\theta\right] - 6.02$$

$$= 21.4 \text{ dBm (AIP to next device)}$$

$$L_{AIP} - dBm = 85.1 \text{ dB (mixer gain)}$$

8.14.6 Power Amplifier

Given:

$L_{AIP} = 21.4$ dBm (output of preceding device),
$E_P = 15.5$ V,
$Z_P = 8\,\Omega$,
$\theta = 0°$.

Find:

$$dB_{amp} = 10\log\left[\frac{E_P{}^2}{0.001Z_P} \times \cos\theta_P\right]$$

$$= 44.8 \text{ dBm (output)}$$

$$dB_{amp} - L_{AIP} = 23.4 \text{ dB (amplifier gain)}$$

$$dB_{amp} - dBm = 108.5 \text{ dB (system gain)}$$

The following equations are reciprocal.
Voltage Amplitude:

$$E_o = \sqrt{(0.001R_s) \times 10^{\left[\frac{\left(\frac{L_{aip}+6.02}{\cos\theta}\right)}{10}\right]}} \qquad (8\text{-}120)$$

Signal Amplitude:

$$L_{aip} = 10\log\left(\frac{E_o{}^2}{0.001R_s}\right)\cos\theta - 6.02 \qquad (8\text{-}121)$$

Note: Voltages are not levels. Voltage amplification is not gain. Level and gain are clearly defined in existing standards. Misuse, even by a majority, is not recognized by the authors.

8.15 Gain Structure Revisited

The earliest patented electrical sound reinforcement system was the product of Western Electric. In the patent application the system was referred to as “Loudspeaking Telephone Outfit.” The foregoing treatment of gain structure in Sections 8.13 and 8.14 is based on the concept of available power. This type of analysis is applicable to all audio systems from Western Electric’s earliest effort to the ones being installed on the day this is being written in 2005. Western Electric’s first patent date for their system occurred in the year 1907. The employment of electrical sound reinforcement systems appears then to be just shy of being a century old.

In the period 1907 to 2005 the technologies involved in the individual elements of a sound reinforcement system have undergone many revolutionary changes from the simplest vacuum electron tube to the latest VLSI (very large-scale integrated circuits). In fact, the physical size of a single early vacuum electron tube was about the same as that of a modern handheld real time analyzer even though the analyzer contains the equivalent of 100,000 or more vacuum tubes. Incidentally, a single vacuum tube wasted more power as heat than the power required to operate a handheld RTA.

In spite of all of the changes in sound system amplifying and signal processing technology, gain structure analysis via the concept of available power is applicable to all. This universality results because this analysis technique does not depend upon whether the system employs matched impedances or not nor does it depend on whether impedances are high or low or mixed values. A tool of such all-encompassing applicability is worthy of being studied and understood. The available power method, however, is not without faults. It is cumbersome to apply in the field because the available power at either the input or output of a general device cannot be determined by a single measurement. One requires at least two measurements, an open circuit voltage measurement with signal present accompanied by an impedance measurement in the absence of signal. We will now explore an alternative system of level analysis, which though restricted in its applicability, is well adapted to the present day electrical properties of the various electronic subsystems from which a sound reinforcement system is constructed.

8.15.1 Two Port Devices

The subsystems constituting an overall sound reinforcement system are members of a class of devices called two port devices. The symbol as well as the significant external variables for an electrical two port such as an amplifier, signal delay, equalizer, transformer, etc. appear in Fig. 8-43.

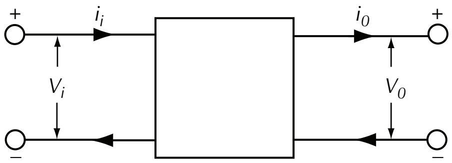


Figure 8-43. An electrical two port device.

The device of Fig. 8-43 features an input port where the quantities of interest are the input voltage V_i and input current i_i. Similarly at the output port the quantities are the output voltage V_0 and output current i_0. The input voltage and current as well as the output voltage and current in general are both time and frequency dependent. At a given frequency, they each may be described by phasors. As drawn, where all four terminals are unique, the device being represented would have a balanced input and a balanced output. If the two lower terminals were each connected to ground, the device being represented would be unbalanced at both input and output. An unbalanced input with balanced output would have only the lower input terminal connected to ground. An unbalanced output with balanced input would have only the lower output terminal connected to ground. In any event, the four quantities are connected by a set of equations the exact form of which depends upon the choice of dependent and independent variables and whether or not the device being described is linear or non-linear in its behavior. In the case of a linear device the describing equations are linear with coefficients whose size do not depend on either voltage or current. Even in a linear system the coefficients can be frequency dependent. In a non-linear device, the coefficients in the describing equation are functions of the various voltages and currents. The coefficients are called the device parameters. These parameters are crucial to the device designer but not of importance here.

8.15.2 Field Adjustment by Voltage Only Analysis

The easiest and most convenient measurement made in the field is that of potential difference between two terminals. Such a measurement should preferably be made by an accurately calibrated wide band scopemeter. Scopemeters combine the functions of oscilloscope, dc volt-ohmmeter, and true rms ac voltmeters. The advantage of such instruments is that one can simultaneously display numeric values of the desired voltage quantity as well as view the signal waveform versus time. By making observations both with test signal applied as well as test signal off it is possible to measure signal voltages as well as determine noise floor voltages and detect if unwanted oscillations are present.

It is necessary at this point to inquire what conditions must be satisfied in order to make voltage only analysis an accurate and viable technique. This can be answered by examining a simplified generic two port along with a source and a load as displayed in Fig. 8-44.

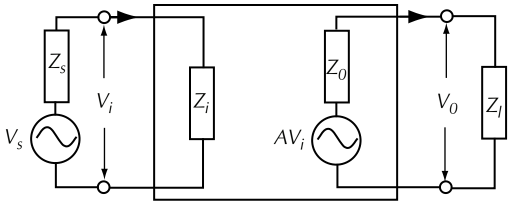


Figure 8-44. Simplified generic two port with signal source and load.

The signal source has a time and frequency dependent open circuit voltage denoted as v_s and an output impedance denoted as Z_s. The input impedance of the two port defined to be v_i/i_i is Z_i. The action of the output port is described by a voltage generator whose open circuit voltage is Av_i in series with an output impedance Z_o. The complex voltage amplification, A, describes the function performed by the two port. If the device being represented were a simple amplifier, A would be a description of its frequency response. If it were an equalizer, A would describe all of the amplitude peaks and dips along with the associated phase changes all as a function of frequency. If the device were a signal delay unit, A would have an all pass character while introducing a negative phase shift proportional to frequency. In sum, the arrangement of Fig. 8-44 can describe any subsystem employed in the overall sound reinforcement system.

What is of importance is the relationship between the source voltage and the voltage that appears across the load connected to the output port. This relationship can be written as

$$\begin{aligned} v_i &= v_s \frac{Z_i}{Z_s + Z_i} \\ v_o &= A v_i \frac{Z_l}{Z_o + Z_l} \\ v_o &= A v_s \frac{Z_i}{Z_s + Z_i} \frac{Z_l}{Z_o + Z_l} \end{aligned} \tag{8-122}$$

Now let's examine the last of Eq. 8-122. Suppose all of the devices composing the system were constructed such that their output impedances were very much smaller than their input impedances. Remember that in a cascade of devices, the output impedance of the first is the source impedance of the following device. Additionally, the input impedance of a following device constitutes the load impedance for the previous device. Then the last equation above could be written

$$\begin{aligned} v_o &= A v_s \frac{1}{\frac{Z_s}{Z_i} + 1} \frac{1}{\frac{Z_o}{Z_l} + 1} \\ &\approx A v_s \end{aligned} \tag{8-123}$$

The justification being that $(Z_s/Z_l) << 1$ and $(Z_o/Z_l) << 1$. Suppose, for example, that the source impedance magnitude whether it is real or complex has an absolute magnitude of the order of 100 Ω while the load impedance similarly has an absolute magnitude of 10,000 Ω or greater. This being the case, the error in each of the approximations above is always less than 1% producing a total error that is always less than 2%. Field voltage measurements themselves are seldom made with such accuracy. When the impedance ratios are satisfied, the magnitude of the voltage amplification expressed in decibels will have at most an inherent error of less than 0.2 dB.

Caution! The effect of interconnecting cable capacitance must be included in calculating the impedance Z_i. To that end, consider Fig. 8-45.

Fig. 8-45 depicts a signal source, the shunt capacitance of the interconnecting cable, and the elements composing the input impedance of a two port device. The cable capacitance C_C can be readily accounted for by incorporating its effect into a calculation of a new value for Z_i. C_C and C_i are in parallel so their values add to produce a total capacitance C_t. Z_i now becomes the parallel combination of C_t and R_i. The magnitude of this impedance is given by

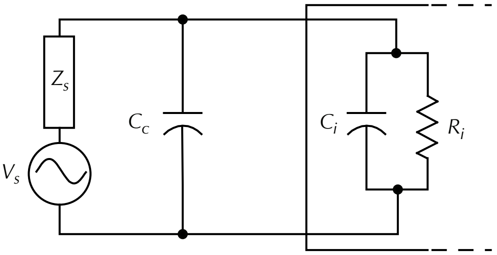


Figure 8-45. Accounting for interconnecting cable capacitance.

$$|Z_i| = \frac{\frac{R_i}{\omega C_t}}{\sqrt{R_i^2 + \left(\frac{1}{\omega C_t}\right)^2}} \tag{8-124}$$

What is needed is an equation that allows us to calculate how much shunt capacitance can be tolerated. Eq. 8-115, when solved for C_t, produces

$$C_t = \frac{\sqrt{R_i^2 - |Z_i|^2}}{\omega R_i |Z_i|} \tag{8-125}$$

It should be clear from Eq. 8-116 that R_i must be greater than the target value of the magnitude of Z_i. Taking 10,000 Ω as the target value, a realistic value for R_i is 15,000 Ω. Additionally we desire the relation to be true throughout the system pass band so ω is taken to be 2π times 20,000 Hz. Upon employing these values C_t becomes 0.59 nF. Even if one allows C_i to be 0.1 nF, a generous value, C_c becomes 0.49 nF. This value would allow several feet of typical interconnecting cable to be employed without the slightest departure from the previously stated error limit. Raising the value of R_i and/or accepting a lower value for the magnitude of Z_i allows for larger values of the total capacitance thus accommodating longer cables. Even when the magnitude of Z_i is allowed to fall to 5000 Ω the error made in the approximation of Eq. 8-114 is still negligible.

The majority, if not all of the subsystems of current manufactures, satisfy the impedance ratio conditions mentioned above including microphones as well as power amplifier-loudspeaker combinations. Microphones present source impedances between about 50 Ω to 300 Ω while amplifiers have inherent output impedances much smaller than the loudspeaker load impedances that they must drive. Now that the voltage analysis only technique has been demonstrated to be a practical approach it will be applied in setting up an example system.

8.15.3 Voltage from Input to Output

Voltage values in volts may be expressed on a logarithmic scale by referencing to a convenient voltage value. There are two reference voltages in common use, one volt and 0.775 V. A voltage of one volt when referenced to one volt and then expressed in decibels becomes 0 dBV. Similarly a voltage of 0.775 V when referenced to 0.775 V and then expressed in decibels becomes 0 dBu. The one-volt reference value is obvious and simple. The 0.775 V reference stems from the fact that the voltage required to produce a power level of 0 dBm in a 600 Ω resistance is 0.775 V. A voltage value of one volt when expressed in dBu is 2.21 dBu. Voltage amplification or attenuation can also be expressed in decibels. Here one is looking at a ratio of two voltages namely output voltage to input voltage. If the voltage ratio is 10, the voltage amplification expressed in decibels is 20 dB. Notice the absence of any third or trailing letter. Suppose the output voltage value is 10 V. This becomes 20 dBV when referenced to one volt. The input voltage must be one volt. This becomes 0 dBV when referenced to one volt. The amplification is the difference between these two values or 20 dB. Similarly the same output voltage in dBu is 22.21 while the input is 2.21 dBu and again the difference is 20 dB. The conclusion is that statements regarding voltage amplification are independent of the choices of voltage reference values.

Mixers are designed so as to minimize noise while maximizing headroom at all points internal to the mixer. This is often accomplished by employing unity amplification at all stages beyond the microphone or line input preamplifiers. The amplifications of these preamplifiers being adjusted as necessary to bring the input signal in question up to the internal operating "zero" value for the mixer in question. Common practice in mixers is to reference signal voltages to 0.775 V with "zero" corresponding to either 0 dBu or +4 dBu on the mixer's metering system. Headroom values in most mixers fall in the range from 10 dB to 20 dB with 20 dB being quite common. Signal to noise ratio is best preserved in an overall system when each subsystem reaches clipping simultaneously. Consider the system of Fig. 8-46.

The mixer features a meter "zero" of +4 dBu, operates with a +15 dB headroom, and clips at +19 dBu. The equalizer operates with 0 dB net voltage amplification, can accept an input of +16 dBu at clipping, and produces an output of +16 dBu at clipping. The amplifier has a voltage amplification of +34.3 dB, is rated at 200 W into 8 Ω, and clips when delivering 400 W into 8 Ω. The upper row of numbers in Fig. 8-46 gives the voltage values expressed in dBu at various points in the system when driven to the clip point. The maximum power output at clipping is 400 W. The lower row of numbers represents normal operating conditions. The system is designed to maintain headroom of +15 dB throughout. In order to accomplish this, two attenuators or voltage sensitivity adjustments are required. As the equalizer can at most see or supply +16 dBu, the output of the mixer must be dropped by 3 dB. If the equalizer's input voltage sensitivity adjustment precedes any internal active circuitry in the equalizer it may be used to accomplish the required attenuation. The same statement can be made with regard to the power amplifier. When external attenuators are required they are placed on the device requiring attenuation. The attenuators themselves when required are designed such that the preceding device always sees the required high impedance.

Although the system has been analyzed with regard to only voltage considerations internally, power must be considered at the system output. Loudspeakers inefficiently convert electric power into acoustic power. The acoustic power is then radiated into space. The electric power converted, the acoustic power radiated, and the distributions of this power in space are all frequency dependent. The questions of the power required and how it should be distributed in space to produce good listening conditions in a given space are subjects that will be treated in detail in the coming chapters.

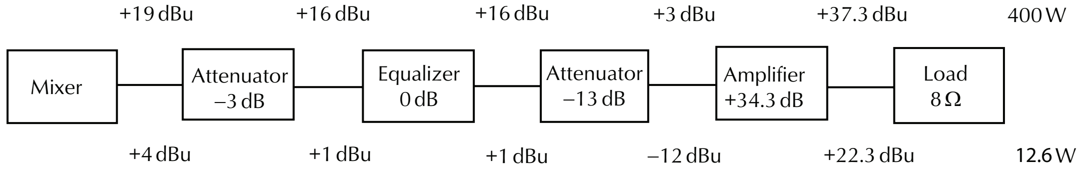


Figure 8-46. Voltage structure of a simple system.

8.15.4 Thermal Noise Levels

Just by being an impedance, a microphone generates thermal noise (*TN*). If no acoustic signal were present, the microphone would still produce a minute output voltage. The open circuit thermal noise relative to 1 V is −198 dB at 1 Hz bandwidth (*BW*) and 1 Ω resistance at a temperature of 293K (68°F). Therefore we can write

$$TN/1\text{V} = -198 + 10\log BW + 10\log R \qquad (8\text{-}126)$$

where,
the bandwidth is in Hz.

Taking our sample microphone with its S_V of −80 dB, we can see that for a bandwidth of 30 Hz–15,000 Hz and a resistance of 150 Ω, we would have:

$$-198 + 10\log(14{,}970) + 10\log 150 = -134\ \text{dBV}$$

The available input power level from this microphone's thermal noise is given by

$$\begin{aligned} L_{AIPN} &= -198 + 10\log(BW) + 30 - 6.02 \\ &= -132\ \text{dBm} \end{aligned}$$

The EIA sensitivity of this microphone is

$$\begin{aligned} G_m &= -80 - 10\log 150 - 50 \\ &= -152\ \text{dBm for an } L_p = 0 \end{aligned}$$

If a talker into the microphone produces a level of $L_p = 74$ dB, then the available talker signal power from the microphone will be

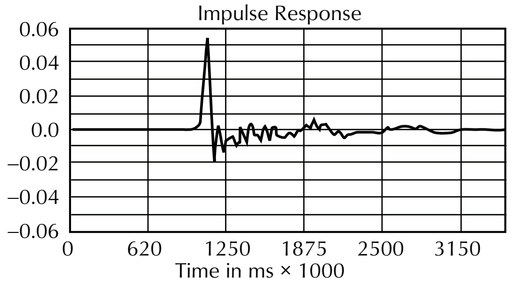
$$\begin{aligned} G_m + L_p &= -152 + 74 \\ &= -78\ \text{dBm} \end{aligned}$$

The *SNR* thus becomes:

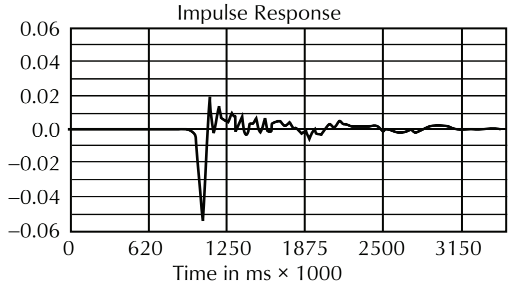
$$\begin{aligned} \textit{Talker level} - \textit{Noise level} &= -78 - 1(-132) \\ &= 54\ \text{dB} \end{aligned}$$

8.15.5 Microphone and Loudspeaker Polarity

It is important to inspect microphone polarity before using multiple microphones in a system. Keep clearly in mind the distinction between polarity and phase and give due consideration to each in the setup of multiple microphone systems. Fig. 8-47 shows impulse responses of in-polarity and out-of-polarity. Loudspeaker polarity needs to be examined as well.



A. In polarity



B. Out of polarity

Figure 8-47. The effects of polarity on impulse response.

8.15.6 Microphone Interconnections

Very low cost PA components may, on occasion, use unbalanced, high-impedance outputs. The microphone cable will in such cases be a single wire with shield cable. If such a microphone must be used in an emergency with a professional mixer, a high-impedance unbalanced to a low-impedance balanced transformer must be used. In the case of a professional low-impedance, balanced output microphone needing to be connected to the high-impedance unbalanced input of a PA component, a reversal of the same transformer will serve the needs.

Microphone interconnections should also be examined to see if phantom powering is present (be sure phantom power is switched off when measuring mixer input impedance).

Low-impedance balanced circuits use two wires with a shield. The overall cable should, of course, be sheathed in a nonconducting jacket. One of the most frequent causes of extraneous noise in a sound system is the microphone-cable combination that is noisy when touched by the user. Be sure to test the microphone-cable combination by rubbing along the cable with your fingers while listening to the sound system set at a high gain.

8.16 Conclusion

The authors take genuine delight in seeing well-designed sound systems drawn in simple line block diagrams with each block labeled as to its gain or loss, input and output impedances (actual and nominal), and link circuit voltages. Such systems are easy to evaluate or to input with an external signal or access a system signal.

Bibliography

G. Chory. "Measuring a Power Amplifier's True Output Impedance," *Syn-Aud-Con Newsletter*, Vol. 1, No. 3 (1974).

D. Davis. "Fundamentals of Audio Transmission," *Syn-Aud-Con Tech Topics*, Vol. 12, No. 6 (1985).

D. Davis and C. Davis. "Electrical Gain and Loss from Voltage and Impedance Measurements," *Syn-Aud-Con Tech Topic,* Vol. 7, No. 3 (1979).

W. J. Kessler. "Vector Impedance Determination with a Simple Electronic Voltmeter," *Syn-Aud-Con Tech Topic*, Vol. 2, No. 4 (1975).

Radio Div., Western Electric Co. Sound Systems: *General Theory and Practice*. Western Electric Co., 1947.

M. C. Sprinkle. *The Gain of Audio Amplifiers*. Page Engineering, Inc., R-1152-0068 (July 1965).

_______. "The Ultimate Noise," *db Magazine* (June 1969)

_______. "How Much Amplifier Power," *Audio* (June 1971).

_______. "The Gain of Audio Amplifiers," *Syn-Aud-Con Tech Topics*, Vol. 7, No. 7 (1980).

Chapter 9

Loudspeaker Directivity and Coverage

by Don Davis

9.1 Essential Definitions 153
  Loudspeaker Coverage ($C_\angle$) 153
  The Acoustic Origin 153
  The Acoustic Center 153
  Directivity Factor (Q) and Coverage Angle ($C_\angle$) 153
  Directivity Plots 155
  Loudspeaker Directivity Factor (Q) 155
  Definition of Q 156
  Geometric Q Transformed to Steradians, sr 156
  Using Steradians for Site Surveys 157
  $C_\angle$, Q, and D_1 for Spherical Segments 158
  Determining Minimum Geometric Q by Loudspeaker Placement 158
9.2 Describing Q More Accurately 158
  A Subtlety Regarding Q by Placement 159
9.3 Relationship Between $C_\angle$ and Q in an Idealized Case 159
9.4 Idealized Loudspeaker Geometry 160
  Classic Method of Obtaining Axial Q 160
  Equal-Angle, Weighted-Area Method 161
  Processing the Polar Plots 162
  The Critical Distance Method 164
  Architectural Mapping 165
  The Dangers 165
9.5 Class D Audio Amplifiers 165
9.6 Sound as a Weapon 166
9.7 An Older View of Q 166
9.8 Summary 167
Bibliography 167

For decades, in the manufacturing of professional sound equipment, the cost of facilities necessary for the accurate measurement of loudspeaker directivity led to closely held data by the manufacturers. Their reasoning was two fold. One, the expense dictated secrecy regarding methods and applications thus withholding from competition. Second, many of them felt that "honest" data was subject to misinterpretation by other than the few experienced users.

Since the first edition of *Sound System Engineering* thirty years ago, a literal revolution in data gathering, data sharing, and skill in data interpretation has occurred worldwide. Once the key parameters had been clearly identified and a cadre of trained users had been developed, the advent of field usable accurate analysis completed the explosion of superior directional devices and their use by competent designers and installers.

9.1 Essential Definitions

9.1.1 Loudspeaker Coverage ($C_\angle$)

Fig. 9-1 illustrates patterns for loudspeakers having the same $C_\angle$ for a given plane but with differing Q. Fig. 9-2 shows patterns for loudspeakers having the same Q but different $C_\angle$. It can be seen that the $C_\angle$ assigned to a given plane of radiation is that angle formed by the –6 dB points (referred to the on-axis reading) and the source center.

The –6 dB point is of interest to the sound engineer inasmuch as that angle should, in the ideal case, intersect the audience area halfway back to the source as compared to where the on-axis beam intersects the audience.

The $C_\angle$ should be specified for as many planes as are felt necessary. The usual case is to specify it for the horizontal and vertical planes.

The new techniques evolving for use in array mapping combined with analysis techniques that tell you with precision the exact distance to the acoustic origin rather than relying on geometry (which may have nothing to do with the real acoustic origin when "off axis"), allow prediction of acoustic performance at the drawing board stage.

9.1.2 The Acoustic Origin

The acoustic origin is the apparent point in space from which the sound emits as measured by the time from the source to the measuring microphone. The physical origin and the acoustic origin do not normally coincide and the acoustic origin is frequency dependent. To obtain signal alignment, knowledge of the acoustic origins is essential.

9.1.3 The Acoustic Center

The acoustic center is found by projecting back to the source from measured $C_\angle$s and finding where they intersect at the source. On occasion, the acoustic center will be in different locations for vertical $C_\angle$s as compared to horizontal $C_\angle$s. This results in a form of acoustic astigmatism that in turn causes a divergence between the array mapping prediction and the measured end result. The acoustic center is normally in front of the driver mounted on a horn and, in many constant directivity devices, it is near the junction of the extended throat and the first cross sectional perspective change.

Accurate mapping of coverage is dependent upon all of these factors:

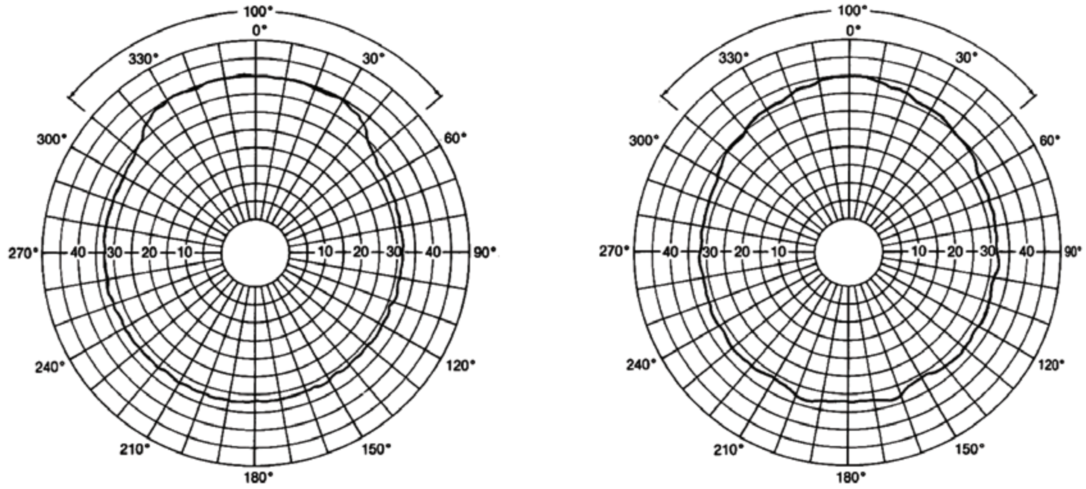
1. Acoustic origin location (signal alignment).
2. Acoustic center (acoustic astigmatism).
3. The angular plot of level changes (making source overlays).
4. The on-axis Q value (for plotting relative Q and changes in $\%AL_{CONS}$).

It is evident to most workers in the field that a great deal of work remains to be done in terms of mapping array coverage, adjusting coverage to its maximum potential, and creatively using mis-coverage for special effects. The tools, for the first time, are available to allow a sufficient number of talented workers at a reasonable cost to foster an explosion of improved concepts, theories, and techniques regarding the directional control of acoustic energy.

9.1.4 Directivity Factor (Q) and Coverage Angle ($C_\angle$)

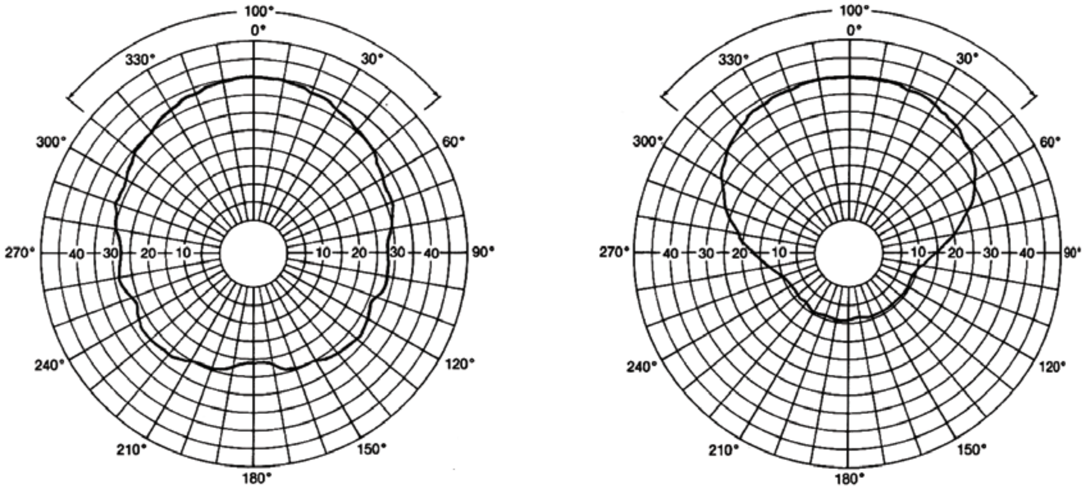
Directivity factor (Q) is always at a point. An area may have an infinite number of points with the same Q, but the area never has a Q. By definition, the average Q around any loudspeaker or array is always unity ($Q = 1.0$).

When loudspeakers are combined in arrays and they are not at wavelengths where mutual coupling operates, the energy from one device may interact with a second device in such a way as to halve one of the $C_\angle$s, such as the vertical in the case of stacked horns. This narrowing of the vertical $C_\angle$ can, on occasion, be highly desirable and help reduce unwanted excitation of a too reverberant room by



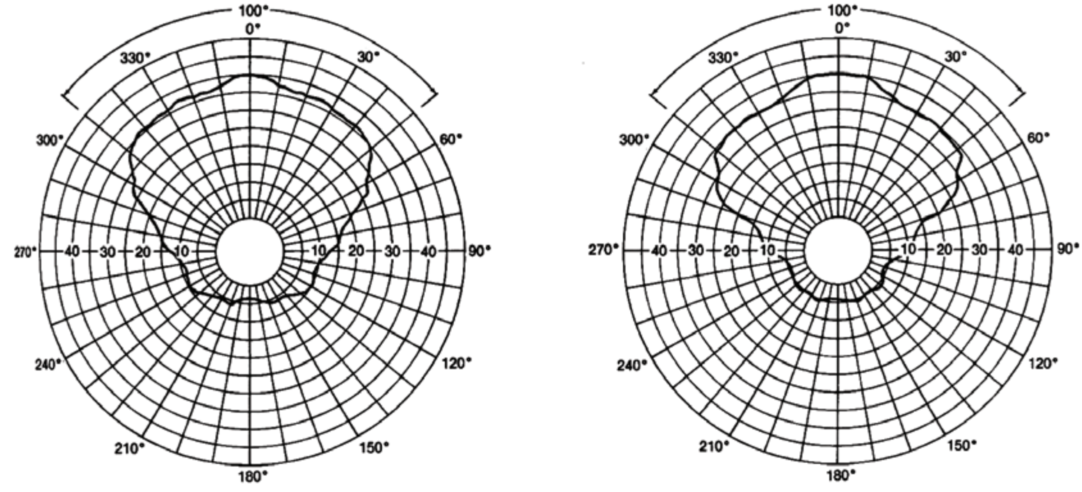
A. Q =3.5

B. Q = 4.9



C. Q = 6.9

D. Q = 8.3



E. Q = 11.4

F. Q = 13.4

Figure 9-1. Polar response charts for loudspeakers possessing a coverage angle of 100° but different directivity factors. (*Courtesy Altec*)

not allowing that energy to fall on the more reflective surfaces but rather confining it to first strike a highly absorptive area. This angular narrowing, however, does not normally increase the Q. While the angle is smaller, all the energy being produced is not passing through the smaller angle, hence, generating more acoustic power per unit area but, rather, is being dissipated in spurious "lobes" generated by the same phase differences that also helped narrow the $C_\angle$. We can talk about woofers in a bin at wavelengths where mutual coupling occurs as a single device.

9.1.5 Directivity Plots

Both Qs and $C_\angle$s are derived from polar response data. This data can take a multitude of forms, but at the present time three techniques are most useful.

Polar Charts. Polar charts look like a polar map of the earth. The transducer being measured is placed on a turntable and rotated in synchronization with the chart revolution in the level recorder.

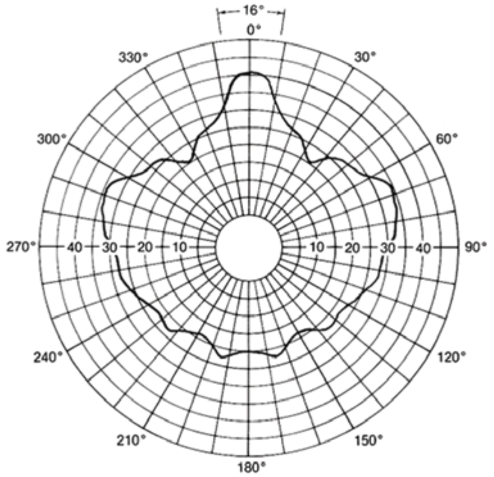
Frequency Charts. Frequency charts are standard frequency response plots done as overlays on the same chart for every 10° or other chosen increments.

The 3-D Plot. The newest techniques are the 3-D plots that allow angular information to be viewed simultaneously with frequency and amplitude information. Fascinating conceptual viewpoints of beaming, lobing, and other aberrations become completely accessible in these plots.

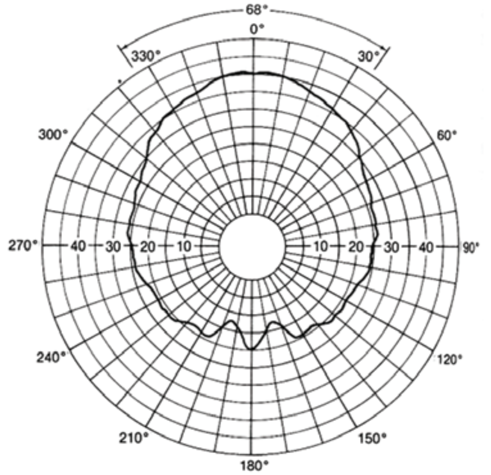
Such uniform level changes allow us to place the farthest listener on-axis (this is hopefully the listener requiring the highest Q) and to orient the device so that listeners closer to the device are on the uniform change in level with angle that exactly compensates for their change in level due to being closer to the source. A loudspeaker that does not change level uniformly with angle would not allow such compensation.

9.1.6 Loudspeaker Directivity Factor (Q)

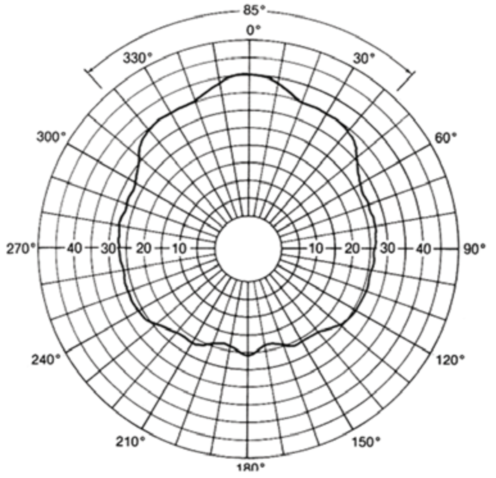
If it were possible to construct a loudspeaker which radiated sound energy only over its $C_\angle$ and nowhere else, then it would be possible to describe its Q from its $C_\angle$. While this is not the case in real-life loudspeakers, and would not really be desirable from the viewpoint of coverage, it is useful to imagine such a case in order to gain a conceptual view of what Q is and how it affects the results we wish to achieve with the loudspeaker.



A. $Q = 11.8$, $C_D = 16°$



B. $Q = 11.1$, $C_D = 68°$



C. $Q = 11.4$, $C_D = 85°$

Figure 9-2. Polar response charts for loudspeakers with similar values of Q but different useful coverage angles. (*Courtesy Altec*).

9.1.7 Definition of Q

First, let us turn to a rigorous, authoritative definition of Q. In Appendix V of the *Handbook of Noise Measurement, Seventh Edition,* by Arnold P. G. Peterson and Ervin E. Gross, Jr.:

> *The directivity factor of a transducer used for sound emission is the ratio of sound pressure squared, at some fixed distance and specified direction, to the mean squared sound pressure at the same distance averaged over all directions from the transducer. The distance must be great enough so that the sound appears to diverge spherically from the effective acoustic center of the source. Unless otherwise specified, the reference direction is understood to be that of maximum response.*

Arnold P. G. Peterson has asked the author to point out that while the author uses various forms of Q to explain room effects, it should be kept in mind that only axial Q is recognized as a property of the loudspeaker by acoustic authorities.

In a note to the above definition, the following additional statement is made:

> *This definition may be extended to cover the case of finite frequency bands whose spectrum may be specified. The average free field response may be obtained, for example, by integration of one or two directional patterns whenever the pattern of the transducer is known to possess adequate symmetry.*

Fortunately, high-quality commercial sound loudspeakers do usually possess adequate symmetry to allow acceptable accuracy with a vertical and horizontal polar response. In cases where the polar response of the loudspeaker is unusually lobed, diagonal polar plots can be taken in addition to the usual horizontal and vertical plots.

9.1.8 Geometric Q Transformed to Steradians, sr

Imagine a perfect spherical radiation pattern from a loudspeaker—and *imagine* is all you can realistically do with real-life loudspeakers. Fig. 9-3 illustrates the spherical surface surrounding such a source. In this case, the average sound pressure, L_P, for the entire surface would equal the L_P measured on any specific axis. Therefore, the directivity ratio would be unity, $Q = 1$.

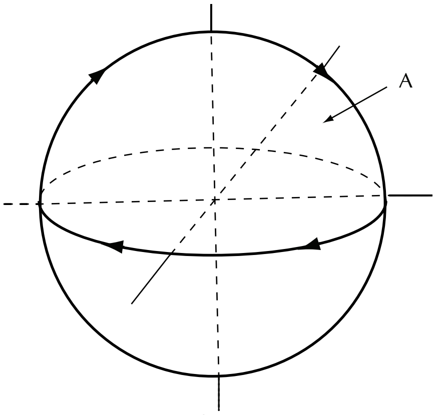


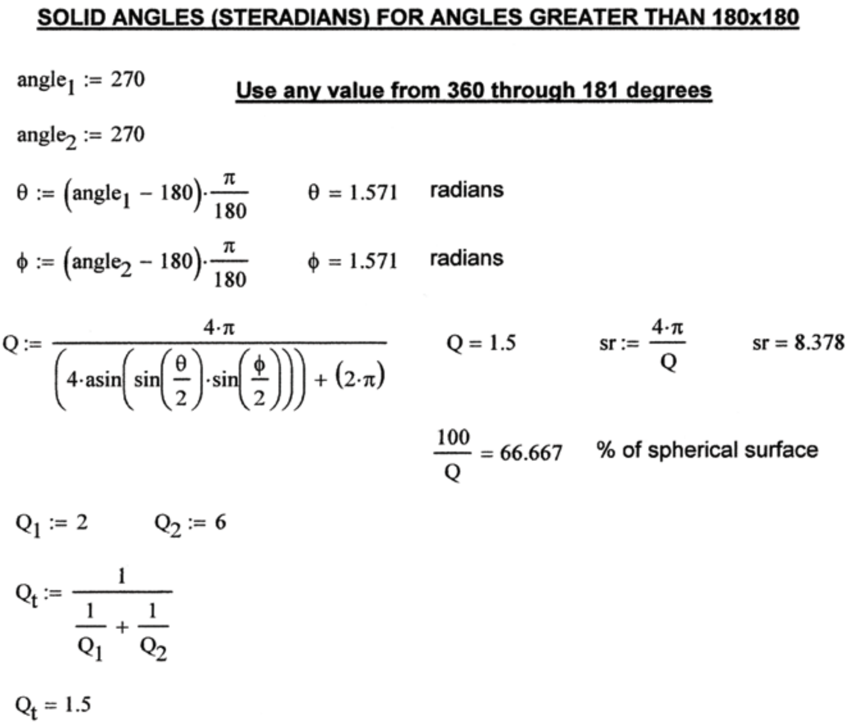
Figure 9-3. Spherical radiation surface.

The next logical question is, "What $C_\angle$ should such a loudspeaker be given?" To compute two angles that represent a sphere we can reason as follows: Traditionally loudspeaker coverage was specified in terms of horizontal and vertical coverage angles in degrees. This was thought of by many as a radius from a point that rotated in a 360° circle horizontally followed by the rotation of the same radius vertically for 360°. In fact, measurements were carried out in exactly this manner. A loudspeaker was rotated on a turntable 360° horizontally and then 360° vertically in front of a fixed microphone position. As a consequence many when asked what would be the coverage angle for a spherical radiator would answer 360° by 360°. Fig. 9-7A uses this labeling method.

The advent of directivity factor Q measurements led to thinking of loudspeaker coverage as area covered. The rotation of a radius of fixed length in a plane through an angle 360° defines all of the points on the circumference of a circle contained in the plane. Now if this circle is rotated about a diameter of the circle through an angle of 180°, the figure so generated is the surface of a sphere whose radius is the same as that of the circle. This is the reason for the rotation limits in spherical coordinates. The azimuthal angle ranges from zero to 360° in the complete description of the spherical surface. It is important to realize that the rotation of a fixed length radius through 360° regardless of the plane in which it occurs describes only a circle.

The equations in Fig. 9-4 do this for all angles for *sr*, Q, and percent of spherical surface covered greater than 180° × 180°. For angles less than 180° × 180°, use Eq. 9-6.

Polar lunes, see Fig. 9-12, can describe angles greater than 180° × 180° as is depicted in Fig. 9-4. It does this by summing steradians. In restudying steradians for this volume, it became apparent that



SOLID ANGLES (STERADIANS) FOR ANGLES GREATER THAN 180x180

$\text{angle}_1 := 270$ **Use any value from 360 through 181 degrees**

$\text{angle}_2 := 270$

$\theta := (\text{angle}_1 - 180)\cdot\frac{\pi}{180}$ $\theta = 1.571$ radians

$\phi := (\text{angle}_2 - 180)\cdot\frac{\pi}{180}$ $\phi = 1.571$ radians

$Q := \frac{4\cdot\pi}{\left(4\cdot\text{asin}\left(\sin\left(\frac{\theta}{2}\right)\cdot\sin\left(\frac{\phi}{2}\right)\right)\right) + (2\cdot\pi)}$ $Q = 1.5$ $sr := \frac{4\cdot\pi}{Q}$ $sr = 8.378$

$\frac{100}{Q} = 66.667$ % of spherical surface

$Q_1 := 2$ $Q_2 := 6$

$Q_t := \frac{1}{\frac{1}{Q_1} + \frac{1}{Q_2}}$

$Q_t = 1.5$

Figure 9-4. MathCAD program depicting great circle coverage.

steradians should have been used instead of Q for Directivity Factor. Both units essentially define area in sound system usage; the audience area is the ideal case. Hopkins-Stryker, critical distance, directivity index, etc., have been re-written for both steradians and Q in this book. When large arrays are involved, the ability to sum all of the individual steradians to obtain the total steradians for the computation of total L_W is invaluable. Additionally, the audience area can be defined as a steradian value

$$sr = \frac{A}{r^2} \tag{9-1}$$

where,
sr is the solid angle in steradians,
A is the audience area,
r is the distance from the source to the listener.

Sr is the SI unit, Q is not. It thus behooves us to become more familiar with this parameter because it is more versatile, internationally accepted, and until now, neglected in audio.

9.1.9 Using Steradians for Site Surveys

When surveying existing sites, such as church auditoriums, gymnasiums, public halls, etc., a superior sales tactic is to provide a demonstration of your capabilities as part of your presentation. Many contractors employ portable hoists allowing quick setups of suitable equipment to match the auditorium's needs.

A quick way to arrive at an accurate estimate of what directivity should be of the loudspeakers to be used is to measure the audience area (in any dimension you wish, feet, meters, inches,) and the radius from that area to a preferred test location. The directivity in Steradians can then be calculated by:

$$sr = \frac{A}{R^2} \tag{9-2}$$

$$Q = \frac{4\pi}{sr} \tag{9-3}$$

where,
A is the area of the audience,
sr is the directivity in Steradians,
R is the radius from the audience area to the desired source location.

We use Steradians because when it becomes apparent that more than one device will be required, the total directivity factor can be determined by simply adding the Steradians of the individual devices and converting the sum to a Q value.

Knowledge of the *total Q* value is then used in the estimation of the intelligibility that will be achieved.

Knowledge of the radius allows levels to be calculated by the inverse square law for the direct sound. From these simple parameters you will have already gained an excellent idea of what equipment will properly fit in the space.

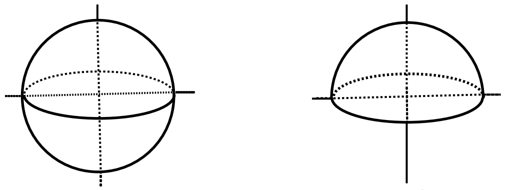
See Table 9-1 for table of directivity factors expressed as Q.

Table 9-1. Loudspeaker Q and Directivity Index D_I

Q	D_I	Q	D_I	Q	D_I
1	0.00	12	10.79	50	16.99
2	3.01	13	11.14	60	17.78
3	4.77	14	11.46	70	18.45
4	6.02	15	11.76	80	19.03
5	6.99	16	12.04	90	19.54
6	7.78	17	12.30	100	20.00
7	8.45	18	12.55	—	—
8	9.03	19	12.79	—	—
9	9.54	20	13.01	—	—
10	10.00	30	14.77	260	24.15
11	10.41	40	16.02		

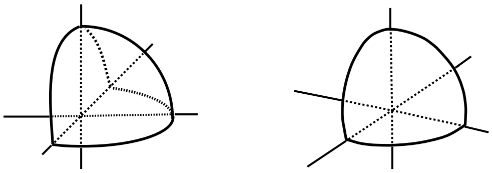
9.1.10 $C_\angle$*, Q, and* D_1 *for Spherical Segments*

Fig. 9-5 depicts a sphere, a hemisphere, a quarter of a sphere, and an eighth of a sphere. Fig. 9-6 also provides additional help in visualizing Q and loudspeaker placement. Always keep in mind, however, that real-life loudspeakers do not follow these simple relationships.



Sphere
Q = 1

Hemisphere
Q = 2



One-quarter sphere
Q = 4

One-eighth sphere
Q = 8

Figure 9-5. Spherical segments.

9.1.11 Determining Minimum Geometric Q by Loudspeaker Placement

	Solid angles	Q
1. Suspended equidistant from all surfaces	4π	1
2. Mounted at center of ceiling or wall surface	2π	2
3. Mounted at intersection of any two surfaces	π	4
4. Mounted at intersection of any three surfaces	π/2	8

$$Q = \frac{4\pi}{sr} \tag{9-4}$$

For $Q = 1.0$ (4π sr), solid angle = 4π.

9.2 Describing Q More Accurately

The measurement of the directivity factor (Q) is always at a point. There can be a series of points within an area that have the same Q thus allowing the concept of an average of Qs within an area. It is normal practice to measure Q on axis (the zero angle axis usually, but not always, being the highest output as well). Let's call this measurement Q_{axis}.

The value Q is both frequency dependent, $Q_{axis(f)}$, and, for real life devices, angularly dependent. Q_{axis} specifies the angle relative to the transducer. For angles other than the "on axis" position we could specify a Q_{rel}.

$$Q_{rel} = Q_{axis}10^{\left(\frac{\pm C_\angle \text{ dB}}{10}\right)} \tag{9-5}$$

where,

$\pm C_\angle$ dB indicates the level in dB of the particular angle relative to the level in dB on axis.

A complete descriptive may be specified by:

$$Q_{rel} = Q_{axis}10^{\left(\frac{\pm C_\angle \text{ dB}}{10}\right)}f \tag{9-6}$$

where,

f is the frequency at which the measurement is made.

Another useful convention would be to agree that where no "f" is specified then the ⅓ octave band at 2000 Hz is indicated.

In the design of a sound system we use:

$$Q_{min(SS)} \tag{9-7}$$

where,

SS stands for single source and which usually is synonymous with Q_{axis} but may, on occasion, actually be a Q_{rel}.

The term "min" indicates that it is the minimum value that will allow the $\%Al_{CONS}$ required at that point.

If more than one source is used, we encounter the term NQ_{min} wherein we increase the Q of the first device proportionately to the number (N) of additional devices (of equal acoustic power output).

We also employ the term Q_{avail} whereby we can calculate the N required for a multiple source system

$$N = \frac{Q_{min}}{Q_{avail}} \tag{9-8}$$

A further refinement is the direct calculation of a distance, D_2, at which the Q_{avail} results in the same ratio of direct-to-reverberant sound as NQ_{min} would have provided.

At the current time we utilize the following Q descriptives: Q_{rel}, Q_{avail}, Q_{axis}, Q_{min} along with the descriptive modifiers: *SS*, $\pm C_{\angle}$, *dB*, *N*, and *f*.

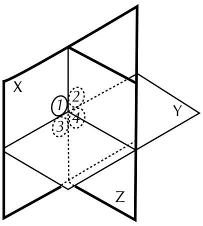
9.2.1 A Subtlety Regarding Q by Placement

An often misinterpreted point with regard to establishing a directivity factor, Q, by placement of the source near a reflecting surface (mirror images) is that the source must be at, not in, the surface, Fig. 9-6A and B.

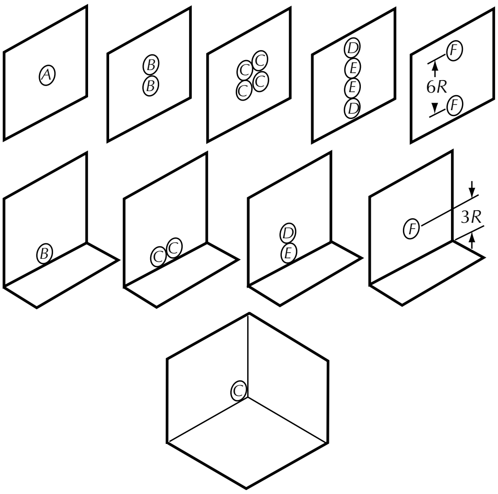
Loudspeakers mounted in the wall will, at lower frequencies, exhibit "mutual coupling" as shown in Fig. 9-6B. When a single speaker is mounted in a wall, half the power goes into another space, if the rear of the loudspeaker is not enclosed. When mounted near the wall, half the power is reflected back into the space.

9.3 Relationship Between $C_{\angle}$ and Q in an Idealized Case

Fig. 9-7 is intended to further assist you in developing a conceptional view of Q and D_I in terms of $C_{\angle}$. Fig. 9-7A shows the angular distribution of a point source defined as the angles formed by the interception of two spherical surface segments. This figure shows what the Q and D_I would be for various combinations of $C_{\angle}$ if all radiation were confined to the angular coverage. While looking at these idealized coverages, remember that real loudspeakers with these same coverage angles also have side, back, top, and bottom lobes that lower the Q, often drastically.



A. Primary images 2, 3, 4 of piston 1 introduced by planes Y and Z.



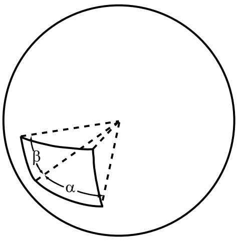
B. Effect of adding pistons and reflecting planes on radiation impedance. All pistons marked with the same letter see the same radiation impedance.

Figure 9-6. Q by placement. (*From Henney's Handbook of Engineering*)

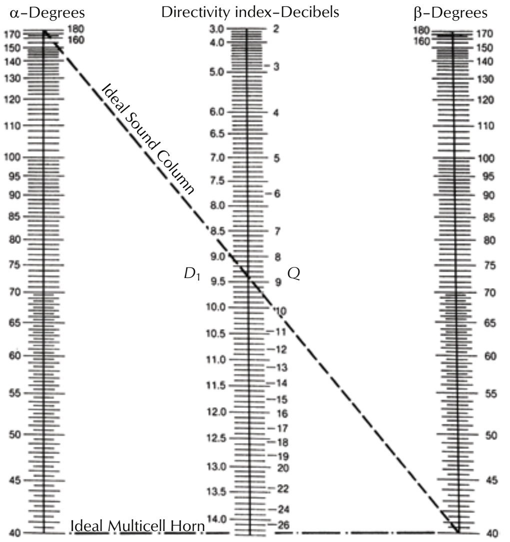
Fig. 9-7B shows that an idealized cone loudspeaker column ($C_{\angle} = 180° \times 40°$) would be limited to a Q of 9. This is because with cones the horizontal coverage would be close to 180°. While the vertical beam narrows, it does not become narrower than 40° anywhere in the useful frequency spectrum. Again, keep in mind that this would be for a perfect unit that had no back, side, top, or bottom radiation.

In contrast, Q would theoretically be 26 for an idealized multicellular horn ($C_{\angle} = 40° \times 40°$). This, again, assumes that no back, side, top, or bottom radiation exists.

Raising Q gives useful on-axis increases of L_P. For example, in the case of the ideal column versus the ideal multicell, a D_I of approximately 14 minus a D_I of 9.5 = 4.5 dB of useful on-axis sensitivity improvement. The multicell would allow the necessary power requirement to be almost ⅓ less than that of the sound column.



A. Angular distribution



B. $C_\angle$, Q, and D_I compared in idealized case

Figure 9-7. Coverage angles, directivity ratio, and directivity index.

9.4 Idealized Loudspeaker Geometry

Loudspeaker directional geometry is of interest to the audio engineer because it allows the development of relative areas associated with different $C_\angle$. The basic formula for finding Q in the idealized case of all energy passing through $C_\angle$ is:

$$Q = \frac{180}{\arcsin\left[\left(\sin\frac{\theta}{2}\right)\times\left(\sin\frac{\phi}{2}\right)\right]} \tag{9-9}$$

Since Q is the inverse of area, we can then write:

$$\text{Relative area} = \frac{\arcsin\left[\left(\sin\frac{\theta}{2}\right)\times\left(\sin\frac{\phi}{2}\right)\right]}{180} \tag{9-10}$$

These geometrical equations are useful in determining the minimum apparent Q that could theoretically be associated with a given requirement of $C_\angle$.

One of the authors was instrumental in the first Q measurements ever published by a commercial sound manufacturer. The methods available today include:

1. The equal-area, multiple-microphone method.
2. The equal-angle, weighted-area method.
3. The critical-distance method.

9.4.1 Classic Method of Obtaining Axial Q

In the noise-measurement field, a relatively standard measurement procedure has been in effect since 1953 (first outlined by Gross and Peterson in the 1953 edition of the *Noise Measurement Handbook*). This method calls for a series of measuring points spaced about the sound source so as to allow each measuring point to represent an equal area on the surface of the sphere. Because of the nature of such geometric patterns, only six such sets of uniformly distributed points are possible. These six sets have 2, 4, 6, 8, 12, and 20 uniformly distributed points. Fig. 9-8 illustrates plane views of such points. The coordinates are given in terms of distances from the center along three mutually perpendicular axes (x, y, z). The "+" refers to the existence of two points, one above the x-y reference plane and one below. When measurements are to be made on a hemisphere, only the four points above the plane are used. Fig. 9-9 shows how such coordinates are utilized to find the desired points.

The length of the vector to the point is found by:

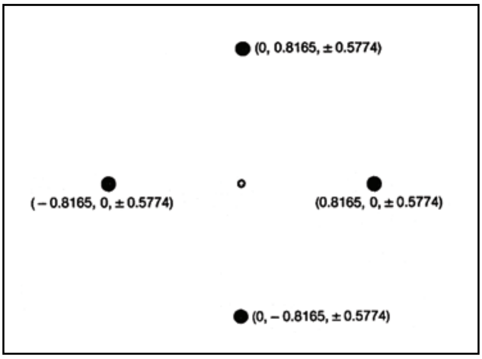
$$\text{Vector length } r = \sqrt{x^2 + y^2 + z^2} \tag{9-11}$$

The angle between the z-axis and the vector is found by:

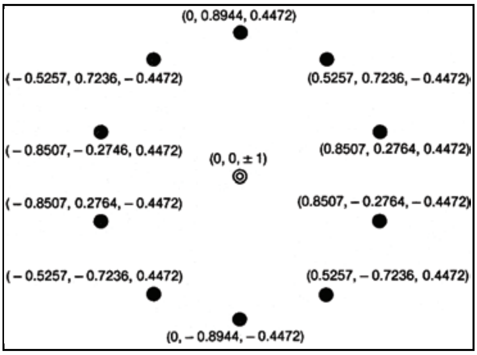
$$\text{Angle} = \operatorname{acos}(z) \tag{9-12}$$

The L_P measured at each of the equal-area points is averaged by converting to power ratios, adding them (dividing them by 2 if only hemispherical measurements are taken, as is often the case), taking 10 times the logarithm of the sum of the powers, and subtracting 10 times the logarithm of the number of points sampled. This gives the average L_P around the sound source being measured; this is identified as $\bar{L}_P$.

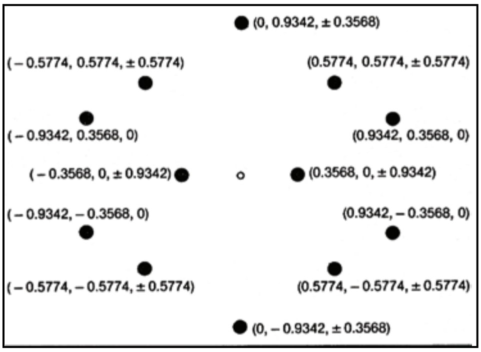
$$\bar{L}_P = 10\log\left(10^{L_{P(1)}/10} + 10^{L_{P(2)}/10} + \ldots + 10^{L_{P(n)}/10}\right) - 10\log(\text{Number of points}) \tag{9-13}$$


(0, 0.8165, ±0.5774)
(−0.8165, 0, ±0.5774)
(0.8165, 0, ±0.5774)
(0, −0.8165, ±0.5774)

A. Eight points


(0, 0.8944, 0.4472)
(−0.5257, 0.7236, −0.4472)
(0.5257, 0.7236, −0.4472)
(−0.8507, −0.2746, 0.4472)
(0, 0, ±1)
(0.8507, 0.2764, 0.4472)
(−0.8507, 0.2764, −0.4472)
(0.8507, −0.2764, −0.4472)
(−0.5257, −0.7236, 0.4472)
(0.5257, −0.7236, 0.4472)
(0, −0.8944, −0.4472)

B. Twelve points


(0, 0.9342, ±0.3568)
(−0.5774, 0.5774, ±0.5774)
(0.5774, 0.5774, ±0.5774)
(−0.9342, 0.3568, 0)
(0.9342, 0.3568, 0)
(−0.3568, 0, ±0.9342)
(0.3568, 0, ±0.9342)
(−0.9342, −0.3568, 0)
(0.9342, −0.3568, 0)
(−0.5774, −0.5774, ±0.5774)
(0.5774, −0.5774, ±0.5774)
(0, −0.9342, ±0.3568)

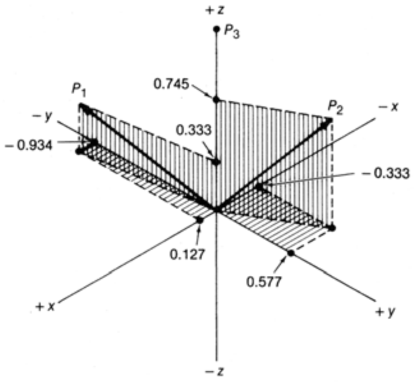
C. Twenty points

Figure 9-8. Plan views of points uniformly distributed on the surface of a sphere of unit radius.

Then Q is found by taking the point of the highest level (*usually* the on-axis point) and subtracting from it the $\bar{L}_P$. The power antilog of this becomes Q.

$$Q = 10^{(L_{P(\text{on axis})} - \bar{L}_P)/10} \tag{9-14}$$

This is called the axial Q. If some other point instead of the on-axis point is chosen, then the calculation becomes the relative Q. This may be further modified into apparent Q by multipliers that will be introduced later.


+z
P_3
0.745
P_1
P_2
−x
−y
0.333
−0.934
−0.333
0.127
0.577
+x
+y
−z

$P_1 = x_1 = 0.127$	$P_2 = x_2 = -0.333$	$P_3 = x_3 = 0$
$y_1 = -0.934$	$y_2 = 0.577$	$y_3 = 0$
$z_1 = 0.333$	$z_2 = 0.745$	$z_3 = 1$
ANGLE = arccos z_1	ANGLE = arccos z_2	ANGLE = arccos z_3
$P_1 = \sqrt{x_1{}^2 + y_1{}^2 + z_1{}^2}$	$P_2 = \sqrt{x_2{}^2 + y_2{}^2 + z_2{}^2}$	$P_3 = \sqrt{x_3{}^2 + y_3{}^2 + z_3{}^2}$

r has been normalized to unity so that
arccos of z is arccos of z/r

Figure 9-9. Locating measuring points on a spherical surface.

Manufacturers of loudspeakers have not used this method but rather have concentrated over the years on gathering polar response data, usually in the horizontal and vertical planes only. Various methods of utilizing such data in order to obtain Q have been tried over the years. While recognizing that the first attempts were crude, it should also be recognized that at the time the cruder methods were used, the alternative was no Q data at all.

9.4.2 Equal-Angle, Weighted-Area Method

Real-life loudspeakers radiate sound over the $C_\angle$ and out of the sides, top, and bottom. In order to find the axial Q, it is necessary to average the *SPL* over the entire space surrounding the source. The method proposed by the author is one derived out of work done by Ben Bauer, C. T. Molloy, and Bob Beavers. This method requires a horizontal and vertical polar plot for each of seven octave bands—125 Hz, 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz, and 8000 Hz.

Fortunately, octave intervals offer more than enough detail to allow accurate planning of the effect of Q on such variables as gain, articulation loss of consonants, etc. Also, most commercial loudspeakers are sufficiently symmetrical in their polar responses to allow the use of a horizontal and vertical polar plot at each octave interval. In some rare cases, additional diagonal plots have to be taken, see Fig. 9-10.

Since both a vertical and a horizontal polar response are taken at each frequency, the manufacturer must then process fourteen polar plots in order to obtain the desired data for a particular loudspeaker.

9.4.3 Processing the Polar Plots

The method of processing the polar plots is illustrated in Table 9-2. Starting at the 0° on-axis point of the polar plot, assign an arbitrary value of 100 dB to the 0° point. Tabulate the relative differences in level, referred to this level, for each 10° point all the way around the horizontal plot. Continue on the

Table 9-2. Weighting Polar Data Taken at 10° Intervals to Correspond to Measurements Taken from Points Surrounded by Equal Surface Areas

Angles	L_P	Rel L_P	Weighting	L_W
0° (on axis)	100	1.000000×10^{10}	0.002418	2.418000×10^{7}
10° & 350°	99	7.943282×10^{9}	0.004730(2)	7.514345×10^{7}
20° & 340°	97	5.011872×10^{9}	0.008955(2)	8.976263×10^{7}
30° & 330°	96	3.981072×10^{9}	0.012387(2)	9.862707×10^{7}
40° & 320°	94	2.511886×10^{9}	0.014990(2)	7.530636×10^{7}
50° & 310°	93	1.995262×10^{9}	0.016868(2)	6.731217×10^{7}
60° & 300°	93	1.995262×10^{9}	0.018166(2)	7.249187×10^{7}
70° & 290°	92	1.584893×10^{9}	0.019007(2)	6.024813×10^{7}
80° & 280°	91	1.258925×10^{9}	0.019478(2)	4.904270×10^{7}
90° & 270°	88	6.309574×10^{8}	0.019630(2)	2.477139×10^{7}
100° & 260°	86	3.981072×10^{8}	0.019478(2)	1.550866×10^{7}
110° & 250°	85	3.162278×10^{8}	0.019007(2)	1.202108×10^{7}
120° & 240°	85	3.162278×10^{8}	0.018166(2)	1.148919×10^{7}
130° & 230°	83	1.995262×10^{8}	0.016868(2)	6.731217×10^{6}
140° & 220°	82	1.584893×10^{8}	0.014990(2)	4.751510×10^{6}
150° & 210°	81	1.258925×10^{8}	0.012387(2)	3.118862×10^{6}
160° & 200°	80	1.000000×10^{8}	0.008955(2)	1.791000×10^{6}
170° & 190°	80	1.000000×10^{8}	0.004730(2)	9.460000×10^{5}
180° (off axis)	80	1.000000×10^{8}	0.002418	2.418000×10^{5}
			Total	**6.934851×10^{8}**

Example Vertical Polar Data at 1000 Hz

Angles	L_P	Rel L_P	Weighting	L_W
10° & 350°	100	1.000000×10^{10}	0.004730(2)	9.460000×10^{7}
20° & 340°	100	1.000000×10^{10}	0.008955(2)	1.791000×10^{8}
30° & 330°	100	1.000000×10^{10}	0.012387(2)	2.477400×10^{8}
40° & 320°	99	7.943282×10^{9}	0.014990(2)	2.381396×10^{8}
50° & 310°	96	3.981072×10^{9}	0.016868(2)	1.343054×10^{8}
60° & 300°	94	2.511886×10^{9}	0.018166(2)	9.126186×10^{7}
70° & 290°	93	1.995262×10^{9}	0.019007(2)	7.584790×10^{7}
80° & 280°	92	1.584893×10^{9}	0.019478(2)	6.174110×10^{7}
90° & 270°	91	1.258925×10^{9}	0.019630(2)	4.942541×10^{7}
100° & 260°	89	7.943282×10^{8}	0.019478(2	3.094385×10^{7}
110° & 250°	87	5.011872×10^{8}	0.019007(2)	1.905213×10^{7}
120° & 240°	85	3.162278×10^{8}	0.018166(2)	1.148919×10^{7}
130° & 230°	79	7.943282×10^{7}	0.016868(2)	2.679746×10^{6}
140° & 220°	75	3.162278×10^{7}	0.014990(2)	9.480509×10^{5}
150° & 210°	72	1.584893×10^{7}	0.012387(2)	3.926414×10^{5}

Table 9-2. (cont.) Weighting Polar Data Taken at 10° Intervals to Correspond to Measurements Taken from Points Surrounded by Equal Surface Areas

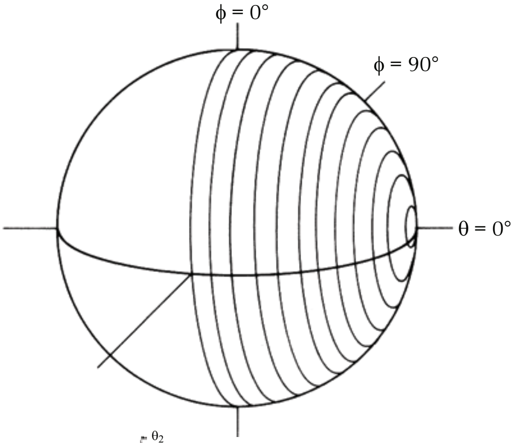
Angles	L_P	Rel L_P	Weighting	L_W
160° & 200°	74	2.511886×10^7	0.008955(2)	4.498789×10^5
170° & 190°	72	1.584893×10^7	0.004730(2)	1.499309×10^5
			Total	$\mathbf{1.238267 \times 10^9}$

$10\log[6.934851 \times 10^8 + 1.238267 \times 10^9] = 92.86\ L_P$

$10^{(100-92.86)/10} = 5.18 = Q$ at 1000 Hz

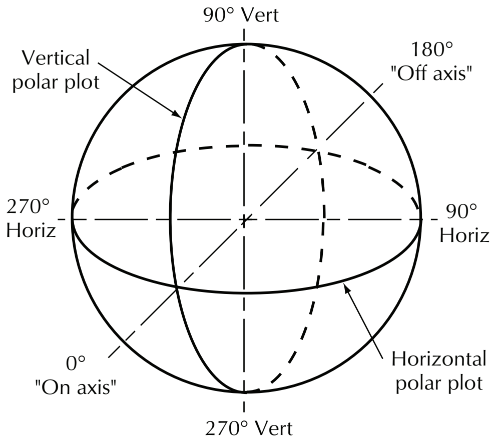
$10\log 5.18 = 7.14\,\text{dB} = D_I$ at 1000 Hz

vertical plot in the same manner but skipping the 0° and 180° points (already recorded). Convert each L_P level to a relative power ratio (*Rel* L_P). Multiply the ratio by a weighting factor proportional to the area surrounding the measuring point in terms of a sphere with a surface of unity. Total all the weighted power ratios (L_{PW}). Then subtract 10 times the logarithm of the sum from the on-axis reading of 100 dB and take the power antilog. This is the axial Q. Figs. 9-11 and 9-12 depict two methods of dividing a sphere into relative areas surrounding each 10° point. Fig. 9-11 shows the zonal method (best for a cone loudspeaker or exactly symmetrical one-cell horns). Fig. 9-12 shows the quadrangle method (best for loudspeaker types other than a single cone or an exactly symmetrical one-cell horn). The dash lines are an extension of the great circles forming the boundaries of the area under consideration. To avoid an overcrowded diagram, only a few such areas are shown in either view. Table 9-3 shows spherical areas.

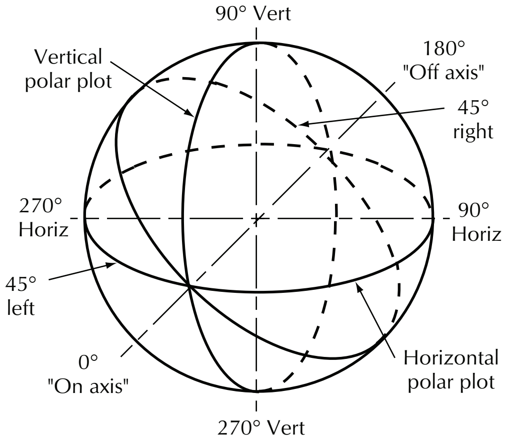


$$\int_{\theta_1}^{\theta_2} \sin\theta d\theta = \cos\theta_1 - \cos\theta_2$$

Figure 9-11. Sphere divided into polar lunes. (*Courtesy The Audio Engineering Society*).



A. Vertical and horizontal only



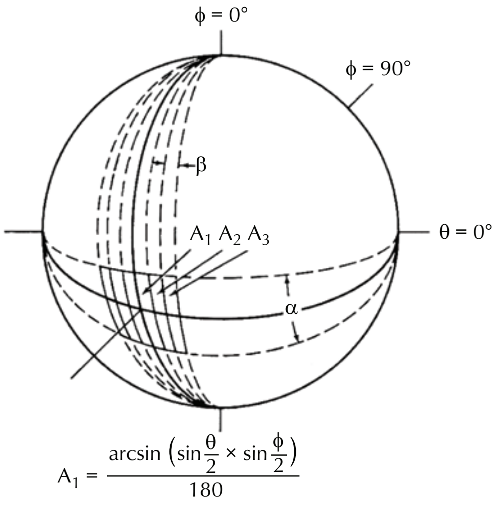
B. Diagonal plots added

Figure 9-10. Sphere divided into zones.

Obviously, the same polar charts used to calculate Q can also be used to obtain $C_\angle$ (the 6 dB-down points from the on-axis reading expressed as an angle). During the calibration of the equipment for the polar responses, the on-axis sensitivity can be measured at the 1 m–1 W, 4 ft–1 W, etc. This should then be translated into the EIA rating, 30 ft at 0.001 W.

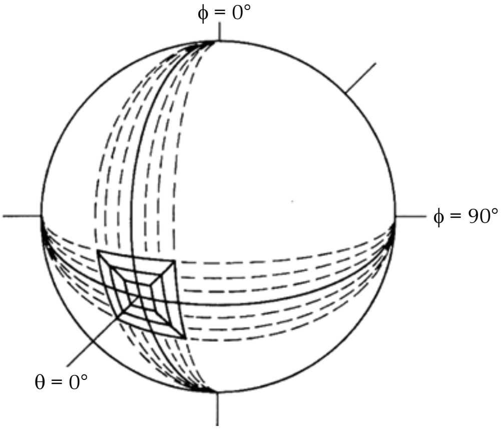
The four primary measurements are:

1. Q in octave bands.
2. $C_\angle$ in octave bands.
3. Axial sensitivity in octave bands.
4. Power handling (program levels).



$$A_1 = \frac{\arcsin\left(\sin\frac{\theta}{2} \times \sin\frac{\phi}{2}\right)}{180}$$

A. Side view



B. Front view

Figure 9-12. Polar response alignments.

Table 9-3. The 10° Spherical Areas for Zones and Quadrangles

10° Intervals	Area of Spherical Zone	Area of Spherical Quadrangle
0° (on axis)	0.003805302	0.004835888
10°	0.030268872	0.037841512
20°	0.059618039	0.071640216
30°	0.087155743	0.099098832
40°	0.112045263	0.119916888
50°	0.133530345	0.134945232
60°	0.150958174	0.145327696
70°	0.163799217	0.152053952
80°	0.171663302	0.155822296
90°*	0.087155744	0.078517492
	Total = 1.000000001	Total = 1.000000004

* The 90° area is to the hemisphere dividing point only. For horizontal and vertical plots there are two on-axis areas, eight areas for angles between 0° and 90°, and four areas at 90°. When right and left diagonal polar plots are added, then there are two on-axis areas, 16 areas for angles between 0° and 90°, and eight areas at 90°.

9.4.4 The Critical Distance Method

The technique most widely used in the actual testing of arrays in large auditoriums and arenas is the measurement of the critical distance, D_c, on the axis of interest. If the engineer has a calibrated Q loudspeaker on hand the measurement of L_D and L_R allows

$$D_c = Ref\ distance\ for\ L_D \times 10^{\frac{(L_D - L_R)}{20}} \tag{9-15}$$

and

$$S\bar{a} = \frac{D_c^{\ 2}}{0.019881Q} \tag{9-16}$$

where,

D_c is the critical distance in ft or m,

Ref distance is a distance in ft or m from the speaker in the free field of the loudspeaker that is ≥10 dB above the reverberant sound field,

L_D is the sound level at the reference distance,

L_R is the sound level in the, hopefully, steady reverberant sound field,

$S\bar{a}$ is the total absorption in ft^2 or m^2 (this will automatically include any modifiers that may be present as well),

Q is the directivity factor of the test loudspeaker at that frequency, octave band, etc.

Having obtained the needed absorption figure, measure the D_c of the array in the same way and use the $S\bar{a}$ obtained from the first measurement to calculate:

$$Q = \frac{D_c^{\ 2}}{0.019881S\bar{a}} \tag{9-17}$$

When measuring D_c, the first measurement is made as far into the reverberant sound field as it is convenient to get—2 to 3 times D_c is ideal.

The second measurement is made by walking toward the loudspeaker (usually on the axis) until the sound level is a minimum of 10 dB higher than the reverberant sound field measurement—again, typically 10 to 15 dB. This insures that the reverberant sound field does not significantly influence the direct sound field measurement. An excellent idea of how steady the reverberant sound field is can be quickly reached during the sampling of it for the reverberant measurement.

The accuracy of this method is defined by precisely the same constraints that apply to the use of Sabine's reverberation equations. This technique should not be used wherever Sabine's reverberation equations cannot be applied.

9.4.5 Architectural Mapping

Since the Pharaoh Zoser and his architect-astronomer-scientist-magician-visier, Imhotep, first built pyramids approximately 6000 years ago (i.e., 3000–4000 B.C.), architectural renderings have changed little. The floor plan and section view are still the architectural mainstays.

Unfortunately, attempting to distort the essentially spherical wavefronts of various kinds as they intersect with linear dimensions on drawings into readily recognizable patterns has, for that same period, defied solution thus leading to Astrolabes', Orrery's, and other spherical devices that allow accurate measurements to be undertaken with visualization of the answer.

Over the centuries, map-makers have exhibited great ingenuity in creating flat maps of spherical areas but always with serious distortion hidden somewhere in the rendering.

The design of the loudspeaker coverage of an audience area has traditionally followed this centuries old pattern to the consequent discomfort of the listener located in one of the hot or dead spots overlooked by these "flat earth" techniques. Today reputable companies provide remarkably useful data and design programs that allow full manipulation of such data in 3-D visual presentations.

9.4.6 The Dangers

It is the temptation to become so involved in obtaining superior coverage that you overlook the very real problems of:

1. When to vary the Q of a device and when you can vary its L_W instead.
2. The cumulative "N" factor.
3. The necessity to use the L_W of only the devices supplying L_D to a point of measurement or observation vs. the total L_W supplying L_R.
4. When to turn from the Peutz equation using Q, V, RT_{60}, etc., to the Peutz equation using L_D, L_R, L_N, and RT_{60}. (This is, in our opinion, one of the most serious flaws in several of the most promoted "flat earth" techniques and one, we fear, not even understood by those advocates.)
5. Solving the $\%AL_{CONS}$ and $PAG = NAG$ before N is accurately determined leads to nonsense such as $\%Al_{CONS}$ predicted from relative dB values obtained from range and device coverage but divorced from the shifting L_W due to N.

You will find in other chapters of this book the constraints placed on how coverage is achieved by the necessity to achieve useful intelligibility, acoustic gain, and the utilization of existing devices. Often beautiful coverage can be achieved by a multiple driver array only to find that the number of sources has reduced intelligibility to an intolerable level, which results in a shortening of the distance that sound can be successfully projected and which, in turn, leads to a whole new approach to coverage—perhaps from high density overhead distribution rather than a multi-driver single source array.

9.5 Class D Audio Amplifiers

Class D audio amplifiers in the multiple kilowatt range can be held in the palm of your hand. A class D amplifier works in very much the same way as a pulse width modulation (PWM) power supply.

Starting with the assumption that the input signal is a standard audio line level signal that is sinusoidal for the frequency ranging from 20 Hz to 20 kHz, this signal is compared with a high-frequency triangle or sawtooth waveform to create the PWM signal. The PWM signal is then used to drive the power stage, creating an amplified digital signal, and finally a low pass filter is applied to the signal to filter out the PWM carrier frequency and retrieve the sinusoidal audio frequency.

Linear amplifiers are inherently very linear in terms of their performance, but are also very inefficient at about 50% typically for a class AB amplifier, whereas a class D amplifier is much more efficient with values in the order of 90% in practical designs. With linear amplifiers, the gain is constant irrespective of bus voltage variations; however with class D amplifiers the gain is proportional to the bus voltage. This means that the power supply rejection ratio (PSRR) of a class D amplifier is zero dB, whereas the PSRR of a linear amplifier is very good. It is common in class D amplifiers to use feedback to compensate for the bus voltage variations.

In linear amplifiers the energy flow was always from supply to the load, and in full bridge class D amplifiers this is also true. In half bridge class D amplifiers the situation is different as the energy flow can be bidirectional which can lead to the "bus pumping" phenomenon. Bus pumping is caused when the bus capacitors are charged up by the energy flow from the load back to the supply; this occurs mainly at the low audio frequencies, i.e., below 100Hz.

At the present time at least two major audio manufacturers are utilizing 4kw multichannel class D amplifiers in conjunction with loudspeaker arrays specifically designed to match the parameters of their amplifiers. Previously, most class D amplifiers have been utilized in automotive sound system applications. Using dedicated DSPs to control the amplitude, phase, and delay to the multiple loudspeakers in the array introduces the possibility of more controlled audience coverage.

9.6 Sound as a Weapon

The frequency of a sound, or the level of a sound, can be used as a weapon. Infra sound is low frequency audio that is below the human range of hearing. Infra sound constantly surrounds us, generated naturally (wind, waves, earthquakes) and by man, building activity, traffic, air conditioners. At higher volumes infra sound of around 7–20Hz can directly affect the human central nervous system causing disorientation, anxiety, panic, nausea, and eventually unconsciousness. The effect can be generated by the extreme low frequencies in church pipe organ music, thereby creating certain emotional effects. (The Broadway play, *Emperor Jones*, used sub-sonic air couplers to cause a feeling of apprehension prior to the Voodoo scenes at the beginning of the 2nd act.) Low-frequency sound generated naturally or by building work in traffic is said to be the cause of reported apparitions and hauntings blamed on the ghostly 19Hz frequency, which matches the resonate frequency of the human eyeball.

There is currently on the market, as I write this book, a long-range acoustic device capable of generating levels at 1m of over 152dB. The LRAD - RX uses directivity and focused acoustic output to clearly transmit critical instructions and warnings well beyond 3000m. Through the use of powerful voice commands and deterrent tones, large safety zones can be created while determining the intent and influencing the behavior of an intruder. Pittsburgh, Pa. police used such a device to disperse protesters gathered outside the G-20 summit held there, the first time such a device had been used on civilians in the United States.

A spokesman for the authorities said:

> *We believe this is highly preferable to the real instances that happen almost every day around the world where officials use guns and other lethal and nonlethal weapons to disperse protesters.*

This device has been successfully deployed against pirates in the Indian Ocean off Somalia by at least one shipping company.

There have been cases reported of entire groups of employees suffering undefined headaches and illnesses until the cause was traced to failure to correctly isolate heavy-duty air conditioning apparatus on the top of the building. Any specification for a sound system must include maximum allowable environmental sound levels, whether from the HVAC system, or from external sources such as nearby railway, airport, or freeway systems.

Carnegie Hall in New York City was built directly above the subway providing a true isolation challenge. In an art museum in Southern California I found the HVAC system mounted in the main exhibit area with no isolation whatsoever, not even a wall.

Sound engineers, utilizing today's very powerful audio amplifiers, and the highly directional loudspeaker arrays, that digital signal processors allow to be constructed, need to make their clients aware of potential hazardous use by unwary operators. Particular care should be exercised in the construction of powerful systems to avoid the generation of high-level impulses from switching circuits and the like.

9.7 An Older View of *Q*

Q, directivity factor, is a valued parameter in the sound engineer's toolbox.

Recently while watching the television news, I found that I was having difficulty hearing the "talking head." My hearing aid was in the bedroom so I just cupped my earlobe. The TV has a calibrated gain control and it was set where I use it with my hearing aid. Much to my surprise I was hearing better with a cupped ear than with a hearing aid on—though it does a good job also. Thinking about what was actually going on, I realized that, yes, cupping my pinna did give me the required gain (about 6 dB) but also removing a good deal of early reflections, allowing me to hear the direct sound more clearly. In fact, the effect of cupping my ears was much like what you hear when comparing one loudspeaker with two that are misaligned. Our living room has large floor to ceiling glass windows as well as a cathedral ceiling and these surfaces were providing sufficiently high level delayed signals to affect speech intelligibility.

As Deward Timothy has demonstrated using PET (Polar Energy Time) measurements, speech intelligibility can be directly influenced not only by gain but also by awkwardly spaced early reflections. We attempt in intelligibility measurements to account for gain, audience coverage, reverberation, when present, and noise.

Years ago while measuring a speech intelligibility problem on the USS Belleau Wood, we found in a narrow walkway above the internal harbor a series of overhead loudspeakers, too-widely spaced, that almost completely destroyed speech intelligibility. A proper solution would have been essentially a line array overhead because of all the steel surrounding surfaces. These were easily the most destructive early reflections we had ever encountered. Experience has taught us awareness of this problem, but there are no easy design criteria for solution at the drawing board stage. I have come to appreciate the effectiveness of the large ear horns used two centuries ago.

9.8 Summary

A general caveat we can give is to use no software in audio or acoustic design work that you have not personally reviewed step-by-step and found to be based on correct basic principles. Computers can generate, at stunning speed and with overwhelming quantity, totally incorrect data when operating from inadequately researched software. Many of the flat earth approaches allow no intuitive safeguards as the mapping employed resembles nothing encountered in real life.

Bibliography

R. J. Bobber. *Underwater Electroacoustic Measurements*. Washington, D.C.: Naval Research Laboratory, 1970.

S. Bridges. "High Q Measurements Confirmed in the Field," *Syn-Aud-Con Tech Topic*, Vol. 3, No. 5 (1975).

P. M. Kending and R. E. Mueser. "Simplified Method for Determining Transducer Directivity Index," *J. Acoust. Soc. Am.*, Vol. 19, No. 4 (July 1947), pp. 691-694.

Wolff and L. Malter. "Directional Radiation of Sound," *J. Acoust. Soc. Am.* (Oct. 1930), pp. 201-241.

Chapter 10

The Acoustic Environment

by Don Davis

10.1 The Acoustic Environment 171
10.2 Dispersion and Diffusion 171
10.3 Inverse Square Law 172
10.4 Atmospheric Absorption 172
10.5 Velocity of Sound 173
10.6 Isothermal vs. Adiabatic 173
10.7 Temperature-Dependent Velocity 174
10.8 The Effect of Altitude on the Velocity of Sound in Air 175
10.9 Typical Wavelengths 175
10.10 Doppler Effect 175
10.11 Reflection and Refraction 176
10.12 Effect of a Space Heater on Flutter Echo 177
10.13 Absorption 177
10.14 Definitions in Acoustics 178
10.15 Classifying Sound Fields 178
10.16 The Acoustic Environment Indoors 181
    The Mean Free Path (*MFP*) 184
    The Build-Up of the Reverberant Sound Field 184
10.17 Conclusion 186
Bibliography 186

10.1 The Acoustic Environment

We are concerned about the effect the acoustic environment has on sound. We need to know the effect of a particular acoustic environment on the unaided talker or musician, on the sound system, if installed, and on unwanted sounds (noise) that may be present in the same environment.

An outdoor environment can often be a "free field." "A sound field is said to be a free field if it is uniform, free from boundaries, and is undisturbed by other sources of sound. In practice it is a field where the effects of the boundaries are negligible over the region of interest." (From the GenRad instruction manual for their precision microphones.)

"Free from boundaries" is the catch phrase here. Anyone who has designed a sound system into a football stadium, a replica of a Greek theater, or a major motor racing course knows first-hand the primary influence of a boundary.

We must also consider:

1. Inverse-square-law level change.
2. Excess attenuation by frequency due to humidity and related factors.

Other factors that can materially affect sound outdoors include:

3. Reflection by and diffraction around solid objects.
4. Refraction and shadow formation by wind and temperature and wind variations.
5. Reflection and absorption by the ground surface itself.

Research in recent years has advanced knowledge of atmospheric absorption significantly from the original base laid by Kneser, Knudsen, followed later by Harris, and more recently by the work of Sutherland, Piercy, Bass, and Evans, Fig. 10-1. This prediction graph is felt to be reliable within +5% for the temperature indicated (20°C) and 10% over a range of 0°C to 40°C.

The June 1977 *Journal of the Acoustical Society of America* had an exceptional tutorial paper entitled "Review of Noise Propagation in the Atmosphere" pages 1403-1418, and included a 96 reference bibliography.

10.2 Dispersion and Diffusion

Dispersion is the process of sorting the rays of light, sound, etc., into their respective wavelengths. Leo Beranek stated:

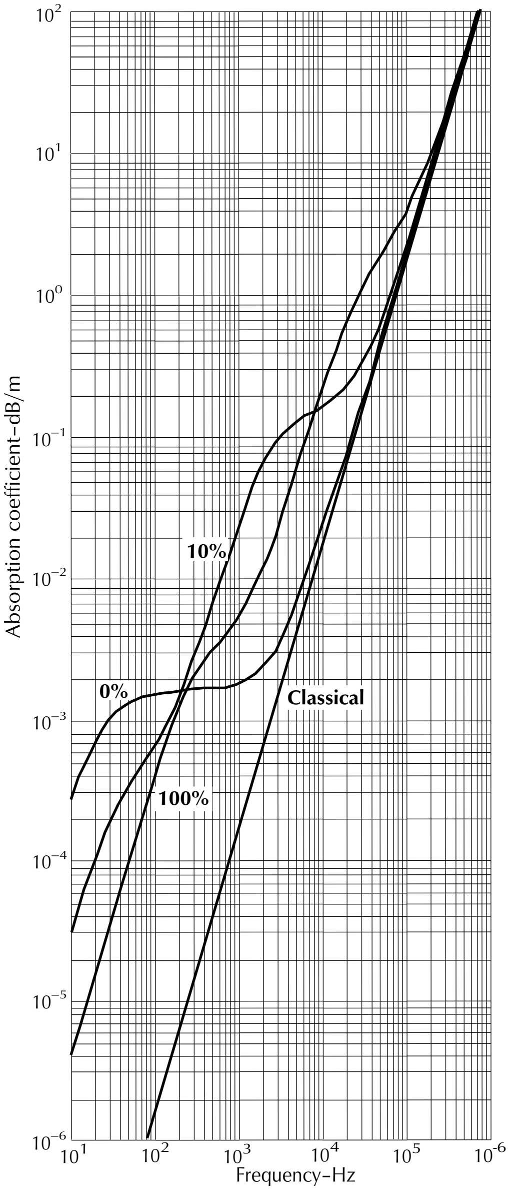


Figure 10-1. Predicted atmospheric absorption in dB/100 m for a pressure of 1 atm, temperature of 20°C, and various values of relative humidity.

Diffusion concerns the spatial orientation of the reverberant sound. Diffusion is considered best when the reverberant sound seems to arrive at the listener's ears from all directions in about equal amounts.

The basic audio parameters are: frequency, wavelength, amplitude, phase (both absolute and relative), polarity, velocity, and time.

Frequency. is related to the sense of pitch in music.

Knowledge of wavelength. is necessary to understand reflection.

Amplitude. is related to loudness.

Absolute phase. is related to time whereas relative phase can be vital to the shape of the wavefront.

Polarity. is of little consequence in a single channel or a truly stereophonic system, but key in a multichannel system.

The velocity of sound. varies with temperature and temperature gradients can change the direction of a wavefront.

Time. is a river that flows endlessly on, but in audio systems latency (signal delay) can cause problems.

10.3 Inverse Square Law

The geometrical spreading of sound from a coherent source (inverse square law rate of level change) which is a change in level of 6 dB for each doubling of distance for a spherical expansion from a point source is well known to most sound technicians.

$$L_P \text{ at measurement point } = \tag{10-1}$$

$$\text{Ref distance} L_P + 20\log\frac{D_r}{D_m}$$

where,
D_r is the reference distance,
D_m is the measured distance.

Not as well recognized is the change in level of 3 dB per doubling of distance for cylindrical expansion from an infinite line source. The ambient noise from a motor race track with the field of cars evenly spread during the early stages of a race can come very close to being effectively an infinite line source.

$$L_P \text{ at measurement point } = \tag{10-2}$$

$$\text{Ref distance} L_P + 10\log\frac{D_r}{D_m}$$

Finally, there is the case of the parallel "loss free" propagation from an infinite area source—the crowd noise viewed from the center of the audience.

Descriptions of the spreading out of sound for coherent sources remains true for incoherent sources as well. The size of the near field may be more restricted and the propagation less directional but the general rate of level change remains the same. Note that this "spreading out" of sound does not constitute absorption or other loss but merely the reduction of power per unit of area as the distance is increased. Unfortunately, other processes also are going on.

10.4 Atmospheric Absorption

These other processes represent actual dissipation of sound energy. Energy is lost due to the combined action of the viscosity and heat conduction of the air and relaxation of behavior in the rotational energy states of the molecules of the air. These losses are independent of the humidity of the air. Additional losses are due to a relaxation of behavior in the vibrational states of the oxygen molecules in the air, because this behavior is strongly dependent on the presence of water molecules in the air (absolute humidity). Both of these energy loss effects cause increased attenuation with increased frequency, Fig. 10-2.

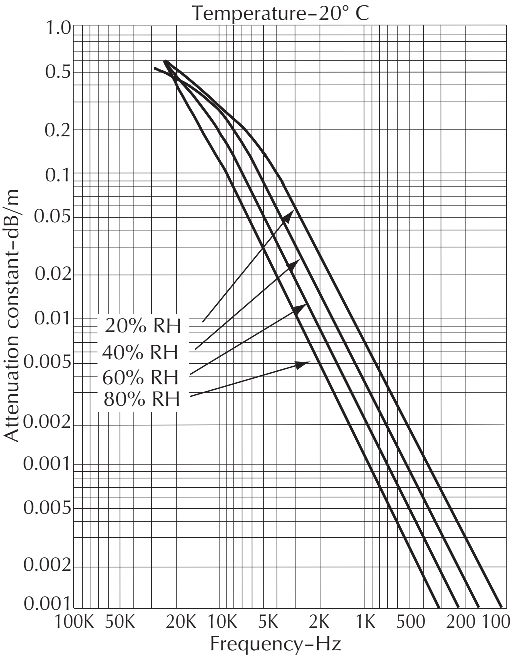


Figure 10-2. Absorption of sound for different frequencies and values of relative humidity.

This frequency-discriminative attenuation is referred to as excess attenuation and must be added to the level change due to divergence of the sound wave. Total level change is the sum of inverse-square-law level change and excess attenuation.

Fig. 10-3 shows the excess attenuation difference between 1000 Hz and 10,000 Hz at various distances.

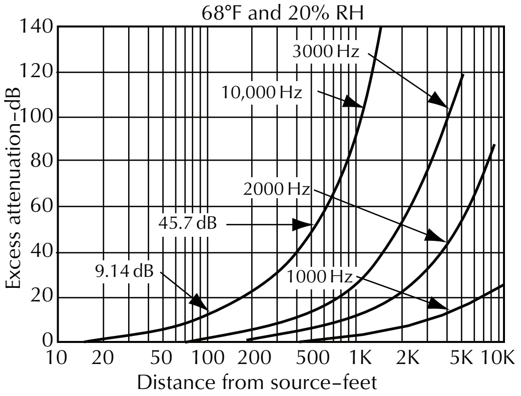


Figure 10-3. Excess attenuation for different frequencies and distances from the source.

10.5 Velocity of Sound

For a given frequency, the relation of the wavelength to the velocity of sound in the medium is

$$\lambda = \frac{c}{f}$$

$$c = \lambda f \qquad (10\text{-}3)$$

$$f = \frac{c}{\lambda}$$

where,

λ is the wavelength in ft or m,

c is the velocity of sound in ft/s or m/s,

f is the frequency in Hz.

In dealing with many acoustic interactions, the wavelength involved is significant and the ability to calculate it is important. Therefore, we need to be able to both calculate and measure the velocity of sound quickly and accurately.

The velocity of sound varies with temperature to a degree sufficient to require our alertness to it. A knowledge of the exact velocity of sound when using signal delayed signal analysis allows very precise distance measurements to be made by observing the frequency interval between comb filters from two sources and then converting from frequency to time and finally to distance.

The velocity of sound under conditions likely to be encountered in connection with architectural acoustic considerations is dependent upon three fundamental factors. These are:

1. Υ is the ratio of specific heats and is 1.402 for diatomic molecules (air molecules).
2. P_S is the equilibrium gas pressure in Newtons per square meter ($1.013 \times 10^5\,\mathrm{N/m^2}$).
3. ρ is the density of air in kilograms per cubic meter ($\mathrm{kg/m^3}$).

$$c = \sqrt{\frac{\gamma P_S}{\rho}} \qquad (10\text{-}4)$$

where,

c is the velocity of sound in m/s.

The density of air varies with temperature and an examination of the basic equations reveals that, indeed, temperature variations are the predominant influence on the velocity of sound in air.

The equation for calculating the density of air is

$$Density\ of\ air = \left[\frac{1.293H}{[1 + 0.00367(°C)](76)}\right] \qquad (10\text{-}5)$$

where,

Density of air is in $\mathrm{kg/m^3}$,

H is the barometric pressure in cm of mercury, Hg,

°C is the temperature in degrees Celsius,

$\frac{9}{5}$ (°C) + 32 = °F,

$\frac{5}{9}$ (°F) – 32 = °C,

Hg in inches times 2.54 equals Hg in centimeters.

10.6 Isothermal vs. Adiabatic

Newton, relying on the simple pressure density relation, came up with the speed of sound as 968 ft/s, which is too low by about 16%.

In the intervening century and a half, Lambert identified no less than nineteen different temperature "scales" that had been proposed by authors of different schemes of thermometry. Among these were the tools that eventually would let the measurement of the change in temperature, caused by rapid compression of the air, to be properly evaluated.

According to Hunt in his truly exceptional book, *The Origin of Acoustics*,

> *The laborious trail from Amontons to Gay–Lussac and Dalton, stretch from the beginning to the end of the eighteenth century, and the prize at its end was merely the fact that the constant in the Towneley–Boyle law is proportional to temperature....*
>
> *As for the raw facts about adiabatic heating and cooling, these were exploited, observed, mis-explained,*

> *inadequately measured, and generally misunderstood, in about that order before enough understanding came to allow them to be used in a final resolution of the sound–speed dilemma.*

John Dalton finally brought this episode to a close with this definitive review of 1802 in which the adiabatic constant for air was defined as 1.4. The rapid temperature changes that occur with increased pressure and decreased pressure in sound, adiabatic, as differentiated from isothermal, where constant temperature is maintained, allowed the calculation of the velocity of sound to match the actual measurements.

Frederick Vinton Hunt's remarkable book is a treasure trove, not only historically, but more importantly for its insight into the multiple side trips and misinterpretations before truly intelligent researchers finally got the message the data at hand contained. Hunt's narrative reveals how publication of experimental data finally congeals in the minds of many, and at the same time, leads to useful progress. The serendipity of research, as well as its science, suggests care in accepting first explanations. As is often the case, reliance on the five physical senses, alone, leads astray.

I recently came across an article "Physical Limits of Computing" by Michael P. Frank, currently at the University of Florida, CISE Department, where he describes the first fully adiabatic central processing unit named, "Pendulum." Mr. Frank was part of a team of graduate students at MIT who designed, outsourced fabrication, and tested the unit. Pendulum was a 12 bit fully adiabatic CPU designed to achieve much lower power requirements than similar devices.

Mr. Frank's paper is outstanding in its outlining of the limits that computing will eventually have to deal with.

Example

If we were to measure a temperature of 72°F and a barometric pressure of 29.92 in, we would first calculate the density of the air according to the data gathered:

$$\frac{5}{9}(72-32) = 22.22\,^\circ\text{C}$$

$$29.92\ \text{in Hg} \times 2.54 = 76\ \text{cm Hg}$$

$$\begin{aligned}\text{Density} &= \frac{1.293(76)}{[1+0.00367(22.22)](76)}\\ &= 1.1955\ \text{kg/m}^3\end{aligned}$$

Having made the metric conversions and obtained the density figure, we can then use the basic equation for velocity

$$\begin{aligned}c &= \sqrt{\frac{1.402(1.013\times 10^5)}{1.1955}}\\ &= 344.67\ \text{m/s}\end{aligned} \tag{10-6}$$

Since we started with the dimensions commonly used here in the United States, we then convert back to them by

$$\frac{344.67\,\text{m}}{1\,\text{s}} \times \frac{100\,\cancel{\text{cm}}}{1\,\cancel{\text{m}}} \times \frac{1.0\,\cancel{\text{in}}}{2.54\,\cancel{\text{cm}}} \times \frac{1\,\text{ft}}{12\,\cancel{\text{in}}} = \frac{1130.81\,\text{ft}}{\text{s}}$$

Typical velocities in other media are shown in Table 10-1.

Table 10-1. Typical Sound Velocities in Various Media (at Approximately 15°C)

Media	Velocity	
	m/s	ft/s
Air	341	1119
Water (Pure)	1440	4724
Water (Sea)	1500	4921
Oxygen	317	1040
Ice	3200	10,499
Marble	3800	12,467
Glass (Soft)	5000	16,404
Glass (Hard)	6000	19,685
Cast Iron	3400	11,155
Steel	5050	16,568
Lead	1200	3937
Copper	3500	11,483
Beryllium	8400	27,559
Aluminum	5200	17,060

10.7 Temperature-Dependent Velocity

The velocity of sound is temperature dependent. The approximate formula for calculating velocity is:

$$c = 49\sqrt{459.4 + {}^\circ\text{F}} \tag{10-7}$$

where,

c is the velocity in feet per second (ft/s),

°F is the temperature in degrees Fahrenheit.

For Celsius temperatures:

$$c = 20.6\sqrt{273 + °C} \tag{10-8}$$

where,
c is the velocity in meters per second (m/s),
°C is the temperature in degrees Celsius.

Therefore, at a normal room temperature of 72.5°F, we can calculate:

$$49\sqrt{459.4 + 72.5} = 1130 \text{ ft/s}$$

10.8 The Effect of Altitude on the Velocity of Sound in Air

The theoretical expression for the speed of sound, c, in an ideal gas (air, for example) is:

$$c = \sqrt{\frac{\gamma P}{\rho}} \tag{10-9}$$

where,
c is the velocity in m/s,
P is the ambient pressure,
ρ is the gas density,
γ is the ratio of the specific heat of the gas at a constant pressure to its heat at constant volume.

Consider the equation

$$PV = RT \tag{10-10}$$

where,
P is the ambient pressure,
V is the volume,
R is the gas constant,
T is the absolute temperature.

Considering the definition of density (ρ), our first equation can be rewritten as:

$$c = \sqrt{\frac{\gamma RT}{M}} \tag{10-11}$$

where,
M is the molecular weight of the gas.

It can be seen that the velocity is dependent only on the type of gas and the temperature and is independent of changes in pressure. This is true because both P and ρ decrease with increasing altitude and the net effect is that atmospheric pressure has only a very slight effect upon sound velocity. Therefore, the speed of sound at the top of a mountain would be the same as at the bottom of the mountain if the temperature is the same at both locations.

10.9 Typical Wavelengths

Some typical wavelengths for midfrequency octave centers are:

Table 10-2. Typical Wavelengths for Mid- frequency Octave Centers

Frequency–Hz	Wavelength–ft
250	4.52
500	2.26
1000	1.13
2000	0.57
4000	0.28
8000	0.14
16,000	0.07

Now suppose the temperature increases 20°F to 92.5°F.

$$49\sqrt{459.4 + 92.5} = 1151 \text{ ft/s}$$

The table of frequencies and wavelengths is shown in Table 10-3.

Suppose we had "tuned" to the peak of a 1000 Hz standing wave in a room first at 72.5°F and then later at 92.5°F. The apparent frequency shift would be:

$$\frac{1151}{1.13} - 1000 = 18.58 \text{ Hz}$$

where,
1151 is the velocity (ft/s) at the temperature of measurement,
1.13 is the wavelength at the original temperature.

Table 10-3. Frequencies and Wavelengths

Frequency (Hz)	Wavelength (ft)
250	4.60
500	2.30
1000	1.15
2000	0.58
4000	0.29
8000	0.14
16,000	0.07

10.10 Doppler Effect

We have all experienced the Doppler effect—hearing the pitch change from a higher frequency to a lower frequency as a train whistle or a car horn comes toward a stationary listener and then recedes into the distance. The frequency heard by the listener due to the velocity of the source, the listener, or some combination of both, is found by:

$$F_L = \left(\frac{c \pm V_L}{c \pm V_S}\right) F_S \qquad (10\text{-}12)$$

where,
F_L is the frequency heard by the listener (observer in Hz),
F_S is the frequency of the sound source in Hz,
c is the velocity of sound in ft/s,
V_L is the velocity of the listener in ft/s,
V_S is the velocity of the sound source in ft/s.

Use minus (–) if V_S in the denominator is coming toward the listener. If the listener, V_L, in the numerator is moving away from the source, use minus (–), and for the listener moving toward the source, use plus (+).

Example

Assume $c = 1130$ ft/s, $V_L = 0$, $V_S = 60$ mi/h (approaching listener), and $f_S = 1000$ Hz

$$\frac{60\,\cancel{\text{mi}}}{1\,\cancel{\text{h}}} \times \frac{1\,\cancel{\text{h}}}{3600\,\text{s}} \times \frac{5280\,\text{ft}}{1\,\cancel{\text{mi}}} = \frac{88\,\text{ft}}{\text{s}}$$

$$F = \left(\frac{1130 - 0}{1130 - 88}\right)1000$$
$$= 1084\text{ Hz}$$

As the sound source passes the listener and recedes, the pitch swings from 1084 Hz to

$$F = \left(\frac{1130 - 0}{1130 + 88}\right)1000$$
$$= 928\text{ Hz}$$

This rapid sweep of 156 Hz is called the Doppler effect. A very large excursion low frequency driver can exhibit Doppler distortion of its signal. Moving vanes in reverberation chambers can produce Doppler effects in the reflected signals that can cause unexpected difficulties in modern spectrum analyzers.

10.11 Reflection and Refraction

Sound can be reflected by hitting an object larger than one-quarter wavelength of the sound. When the object is one-quarter wavelength or slightly smaller, it also causes diffraction of the sound (bending around the object). Refraction occurs when the sound passes from one medium to another (from air to glass to air, for example, or when it passes through layers of air having different temperatures). The velocity of sound increases with increasing temperature. Therefore, sound emitted from a source located on the frozen surface of a large lake on a sunny day will encounter warmer temperatures as the wave diverges upward, causing the upper part of the wave to travel faster than the part of the wave near the surface. This causes a lens-like action to occur which bends the sound back down toward the surface of the lake, Fig. 10-4.

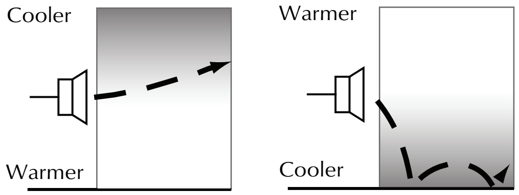


Figure 10-4. Effect of temperature differences between the ground and the air on the propagation of sound.

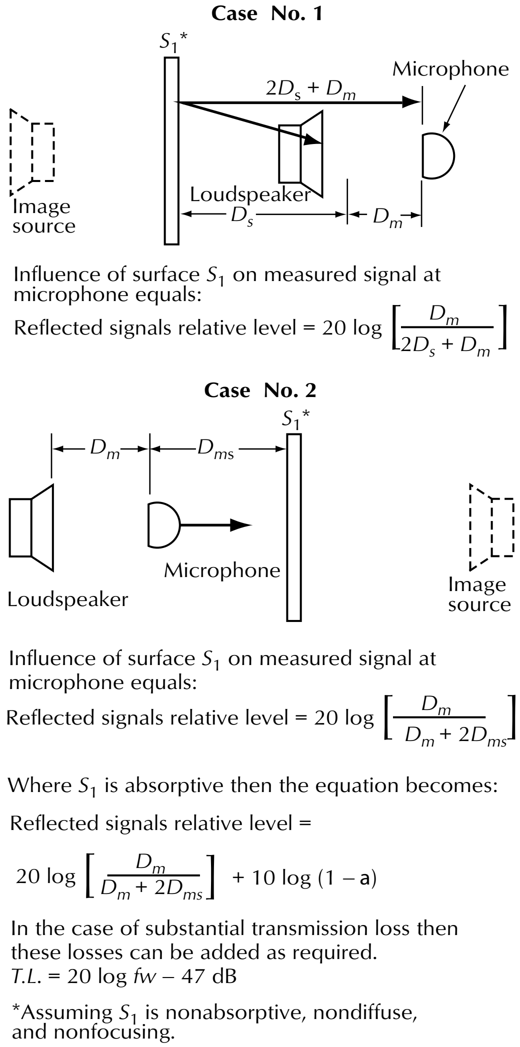
Sound will travel great distances over frozen surfaces on a quiet day. Wind blowing against a sound source causes temperature gradients near the ground surface that result in the sound being refracted upward. Wind blowing in the same direction as the sound produces temperature gradients along the ground surface that tend to refract the sound downward. We hear it said, “The wind blew the sound away.” That is not so; it refracted away. Even a 50 mph wind (and that’s a strong wind) cannot blow away something traveling 1130 ft/s:

$$\frac{1130\text{ ft}}{1\text{ s}} \times \frac{3600\text{ s}}{1\text{ h}} \times \frac{1\text{ m}}{5280\text{ ft}} = 770.45\text{ mi/h}$$

770.45 mi/h is the velocity of sound at sea level at 72.5°F.

Wind velocities that vary with elevation can also cause “bending” of the sound velocity plus or minus the wind velocity at each elevation.

Reflections from large boundaries, when delayed in time relative to the direct sound, can be highly destructive of speech intelligibility. It is important to remember, however, that a reflection within a nondestructive time interval can be extremely useful. Reflections that are at or near (within 10 dB) equal amplitude and that are delayed more than 50 ms require careful attention on the part of a sound system designer. Fig. 10-5 shows how to calculate probable levels from a reflection. Fig. 10-6 shows other influences. Calculation of the time interval is found by:



Influence of surface S_1 on measured signal at microphone equals:

Reflected signals relative level = $20 \log \left[\frac{D_m}{2D_s + D_m}\right]$

Influence of surface S_1 on measured signal at microphone equals:

Reflected signals relative level = $20 \log \left[\frac{D_m}{D_m + 2D_{ms}}\right]$

Where S_1 is absorptive then the equation becomes:

Reflected signals relative level =

$$20 \log \left[\frac{D_m}{D_m + 2D_{ms}}\right] + 10 \log (1 - a)$$

In the case of substantial transmission loss then these losses can be added as required.
T.L. = 20 log *fw* – 47 dB

*Assuming S_1 is nonabsorptive, nondiffuse, and nonfocusing.

Figure 10-5. Calculating relative levels of reflections.

$$\frac{1000}{c}(D_R - D_D) = \textit{Time interval} \text{ (in ms)} \quad (10\text{-}13)$$

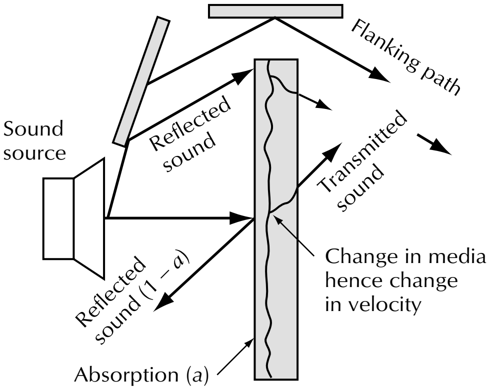
where,

c is the velocity of sound in ft/s or m/s,

D_R is the distance in ft or m traveled by the reflection,

D_D is the distance the direct sound traveled in ft or m.

A large motor speedway used to make very effective use of ground reflections on the coverage of the grandstands behind the pit area. The very high temperature gradients encountered warp the sound upward during the hot part of the day and in the cool of the morning, the ground reflection helps with the coverage of the near seating area. The directional devices are aimed straight ahead along the ground rather than up at an angle and when the temperature gradient "bends" the sound upward, it's still covering the audience area effectively, Fig. 10-4.



Mass Law

T.L. = 20 log [fw] – 47 dB
f = frequency in Hz
w = weight of barrier in kg/m^2
$a = 1 - 10^{(-\text{dB}/10)}$
dB = 10 log (1 – a)

Figure 10-6. Absorption, reflection, and transmission of boundary surface areas.

One caution about using ground reflections in northern climes. A heavy snow fall can provide unbelievable attenuation as the authors can attest after trying to demonstrate, years ago, a high level sound system the day after a blizzard in Minnesota.

10.12 Effect of a Space Heater on Flutter Echo

Velocity of sound increases with an increase in temperature; therefore, the effect of an increase in temperature with an increase in height is a downward bending of the sound path. This illustrates why feedback modes change as air conditioners, heating, or crowds dramatically change the temperature of a room, Fig. 10-7.

10.13 Absorption

Absorption is the inverse of reflection. When sound strikes a large surface, part of it is reflected and part of it is absorbed. For a given material, the absorption coefficient, (a) is:

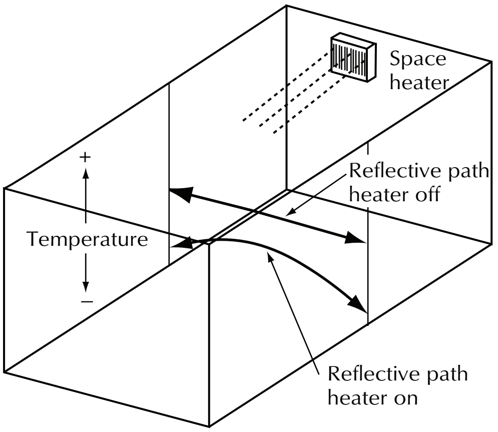


Figure 10-7. Effect of thermal gradients in a room.

$$a = \frac{E_A}{E_I} \qquad (10\text{-}14)$$

where,

E_A is the absorbed acoustic energy,

E_I is the total incident acoustic energy (i.e., the total sound),

$(1 - a)$ is the reflected sound.

This theoretically makes the absorption coefficient some value between 0 and 1. For $a = 0$, no sound is absorbed; it is all reflected. If a material has an a of 0.25, it will absorb 25% of all sound energy having the same frequency as the absorption coefficient rating, and it will reflect 75% of the sound energy having that frequency.

Example

An anechoic room absorbs 99% of the energy received from the sound source. What percentage of the L_P from the source is reflected? Assume 10W of total energy output from the source. Then the chamber absorbs 9.9W of it.

$$10\log\frac{10\text{ W}}{0.1\text{ W}} = 20\text{ dB}$$

Therefore, the L_P drops by 20dB also

$$100 \times 10^{-\text{dB}/20} = 10\%\text{ reflected }L_P$$

In other words, 10% of the L_P returns as a reflection. If the sound source had directed an L_P of 100dB signal at the wall of the chamber, a signal of 80dB would be reflected back. Remembering how dB are combined, we can see that this reflection will not change the 100dB reading of the direct sound by a discernible amount on any normal sound level meter.

The desirability of a reflective surface can be seen when it is realized that the direct sound and the reflected sound from a single surface can combine to be as much as 3dB higher than the direct sound alone. If the loudspeakers are directed to reflect off the ground during the cool early morning hours; then when the refraction effect of the sun on the hard surfaces causes the sound to bend upward during the hot part of the day, the sound bends up into the grandstand area. Most of the time, the reflected sound is assisting the direct sound, thereby saving audio power.

10.14 Definitions in Acoustics

Sound Energy Density—is the sound per unit volume measured in joules per cubic meter.

Sound Energy Flux—is the average rate of flow of sound energy through any specified area. The unit is joules per second (joules per second are called watts).

The Sound Intensity—(or sound energy flux density) in a specified direction at a point is the sound energy transmitted per second in the specified direction through unit area normal to this direction at the point. The unit is watts per square meter.

Sound Pressure—is exerted by sound waves on any surface area. It is measured in Newtons per square meter (now called pascals). The sound pressure is proportional to the square root of the sound density.

The Sound Pressure Level—(in decibels of a sound)—20 times the logarithm to the base 10 of the ratio of the pressure of this sound to the reference pressure. Unless otherwise specified, the reference pressure is understood to be 0.00002 N/m^2 (20micropascals or 20µPa).

The Velocity Level—(in decibels of a sound) 20 times the logarithm to the base 10 of the ratio of the particle velocity of the sound to the reference particle velocity. Unless otherwise specified, the reference particle velocity is understood to be 50×10^{-9} meters per second (m/s).

The Intensity Level—(in decibels of a sound) 10 times the logarithm to the base 10 of the ratio of the intensity of this sound to the reference intensity. Unless otherwise specified, the reference intensity shall be 10^{-12} watts per square meter (W/m^2).

10.15 Classifying Sound Fields

Free Fields. A sound field is said to be a free field if it is uniform, free of boundaries, and is undisturbed by other sources of sound. In practice, it is a field in which the effects of the boundaries are negligible over the region of interest. The flow of sound energy is in one direction only. Anechoic chambers and well-above-the-ground outdoors are free fields. The direct sound level from a sound source in a free field is labeled L_D.

Diffuse (Reverberant) Fields. A diffuse or reverberant sound field is one in which the time average of the mean square sound pressure is everywhere the same and the flow of energy in all directions is equally probable. This requires an enclosed space with essentially no acoustic absorption. The reverberant sound level is labeled L_R.

Semireverberant Fields. A semireverberant field is one in which sound energy is both reflected and absorbed. The flow of energy is in more than one direction. Much of the energy is truly from a diffused field; however, there are components of the field that have a definable direction of propagation from the noise source. The semireverberant field is the one encountered in the majority of architectural acoustic environments. The early reflections, i.e., under 50 ms after L_D, are labeled L_{RE}.

Pressure Fields. A pressure field is one in which the instantaneous pressure is everywhere uniform. There is no direction of propagation. The pressure field exists primarily in cavities, commonly called couplers, where the maximum dimension of the cavity is less than ⅙ of the wavelength of the sound. Because of the ease of repeatability, this type of measurement is used by the National Bureau of Standards, NBS, when they calibrate microphones. At low frequencies the pressure field can be large, i.e., big enough for a listener to sit in.

Ambient Noise Field. The ambient noise field is comprised of those sound sources not contributing to the desired L_D, (i.e., active sources). The ambient noise level is labeled L_N.

Outdoor Acoustics. If, for example, the ambient noise level measured 70 dBA (a not unreasonable reading outdoors) and the most *SPL* you could generate at 4 ft was 110 dB L_P how far could you reach before your signal was submerged in noise?

$$\begin{aligned} 110L_P - 70L_P &= 40\text{ dB} \\ 20\log\frac{x}{4} &= 40\text{ dB} \\ x &= 4 \times 10^{40/20} \\ &= 400\text{ ft} \end{aligned}$$

The problem actually is more complicated than this outdoors, but this serves as an illustration of how to begin.

We have now touched on the most important basics of the acoustics environment outdoors. Before going indoors, let us apply some of this knowledge to a series of ancient outdoor problems. A simple rule of thumb dictates that when a change of +10 dB occurs, the higher level will be subjectively judged as approximately twice as loud as the level 10 dB below it. While the computation of loudness is more complex than this, the rule is useful for midrange sounds. Using such a rule, we could examine a sound source radiating hemispherically due to the presence of the surface of the earth. Fig. 10-8 shows sound in an open field with no wind. The sound at 100 ft is one-half as loud as that at 30 ft, although the amplitude of the vibration of the air particles is roughly one-third. Similarly, the sound at 30 ft is one-half as loud as the sound at 10 ft. Because the sound is outdoors, atmospheric effects, ambient noise, etc., cause difficulty for the talker and listener. The ancients learned to place a back wall behind the talker, and many Native American council sites were at the foot of a stone cliff so the talker could address more of the tribe at one time. Fig. 10-9 illustrates how a reflecting structure can double the loudness as compared to the totally open space. The weather and some noise still interfere with listening.

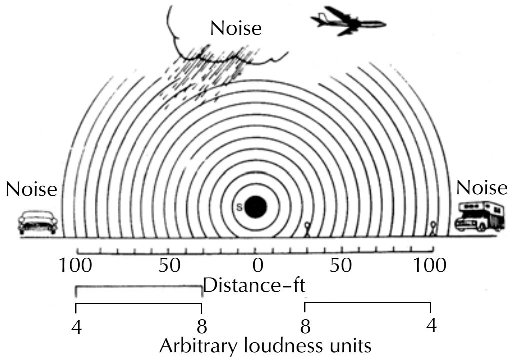


Figure 10-8. Sound in an open field with no wind.

Fig. 10-10 illustrates the absorptive effect of an audience on the sound traveling to the farthest

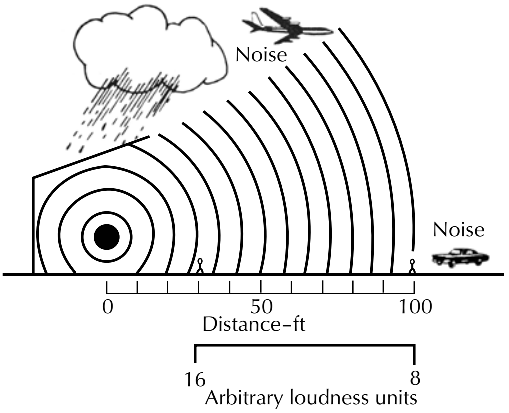


Figure 10-9. Sound from an orchestra enclosure in an open field with no wind.

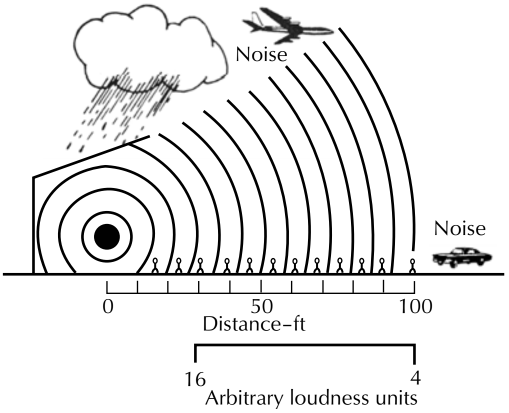
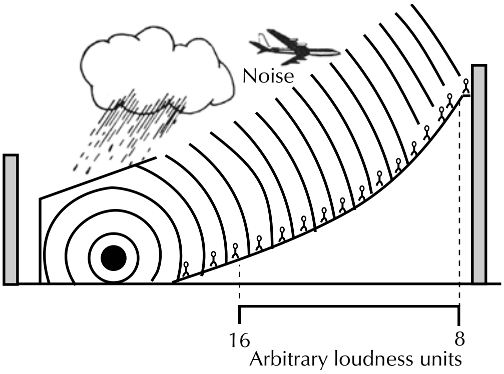


Figure 10-10. Sound from an orchestra enclosure with an audience.

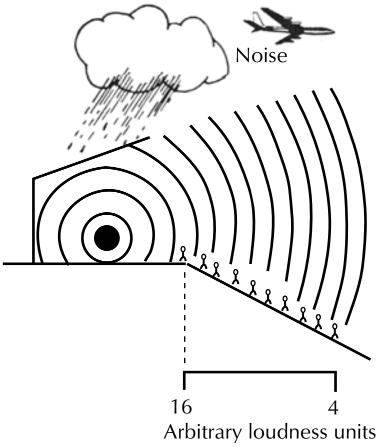
listener. Fig. 10-11 shows the right way and the wrong way to arrange a sound source on a hill. In Fig. 10-11A the loudness of the sound at the rear of the audience is enhanced by sloping the seating upward. In addition, the noise from sources on the ground is reduced. Fig. 10-11B is a poor way to listen outdoors. The sound at the rear is one-half as loud as it is at the rear in Fig. 10-11A.

While the Bible doesn't say which way Jesus addressed the multitudes, we can deduce from the acoustical clues present in the Bible text that the multitude arranged themselves above him because:

1. He addressed groups as large as 5000. This required a very favorable position relative to the audience and a very low ambient noise level.
2. Upon departing from such sessions, He could often step into a boat in the lake, suggesting He was at the bottom of a hill or mountain.



A. Correct way



B. Wrong way

Figure 10-11. Sound sources and audiences on a hill.

We can further surmise that the reason Jesus led these multitudes into the countryside was to avoid the higher noise levels present even in small country villages.

The Greeks built their amphitheaters to take advantage of these acoustical facts:

1. They provided a back reflector for the performer.
2. They increased the talker's acoustic output by building megaphones into the special face masks they held in front of their faces to portray various emotions.
3. They sloped the audiences upward and around the talker at an included angle of approximately 120° realizing, as many modern designers do not seem to, that man does not talk out of the back of his head.
4. They defocused the reflective "slapback" by changing the radius at the edges of the seating area.

Because there were no aircraft, cars, motorcycles, air conditioners, etc., the ambient noise levels were relatively low, and large audiences were able to enjoy the performances. They had discovered absorption and used jars partially filled with ashes (as tuned Helmholtz resonators) to reduce the return echo of the curved stepped seats back to the performers. It remained only for some unnamed innovative genius to provide walls and a roof to have the first auditorium, "a place to hear," Fig. 10-12. No enhancement of sound is provided in Fig. 10-12 because there is no reverberation in a room whose walls are highly sound absorbent.

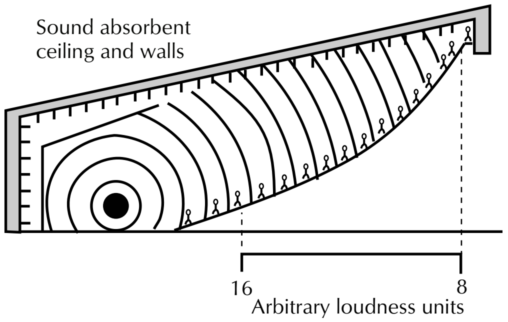


Figure 10-12. Means of eliminating noise and weather while preserving outdoor conditions.

Sometimes acoustic progress was backward. For example, the Romans, when adopting Christianity, took over the ancient echo ridden pagan temples and had to convert the spoken service into a chanted or sung service pitched to the predominant room modes of these large, hard structures. Today, churches still often have serious acoustical shortcomings and require a very carefully designed sound system in order to allow the normally spoken word to be understood.

It is also of real interest to note that in large halls and arenas the correct place for the loudspeaker system is most often where the roof should have gone if the building had been designed specifically for hearing. A loudspeaker is therefore usually an electroacoustic replacement for a natural reflecting surface that has not been provided.

10.16 The Acoustic Environment Indoors

The moment we enclose the sound source, we greatly complicate the transmission of its output. We have considered one extreme when we put the sound source in a well-elevated position and observed the sound being totally absorbed by the "space" around it. Now, let us go to the opposite extreme and imagine an enclosed space that is completely reflective. The sound source would put out sound energy, and none of it would be absorbed. If we continued to put energy into the enclosure long enough, we could theoretically arrive at a pressure that would be explosive. Human speech power is quite small. It has been stated by Harvey Fletcher in his book *Speech and Hearing in Communication* that it would take "…500 people talking continuously for one year to produce enough energy to heat a cup of tea." Measured at 39.37 in (3.28 ft), a typical male talker generates 67.2 dB-SPL, or 34 microwatts (µW) of power, and a typical female talker generates 64.2 dB-SPL, or 18 µW. From a shout at this distance (3.28 ft) to a whisper, the dB L_P ranges from 86 dB to 26 dB, or a dynamic range of about 60 dB. Not only does the produced sound energy tend to remain in the enclosure (dying out slowly), but it tends to travel about in the process.

Let us now examine the essential parameters of a typical room to see what does happen. First, an enclosed space has an internal volume (V), usually measured in cubic feet. Second, it has a total boundary surface area (S), measured in square feet (ft^2) (floor, ceiling, two side walls, and two end walls). Next, each of the many individual surface areas has an absorption coefficient. The average absorption coefficient (a) for all the surfaces together is found by

$$\bar{a} = \frac{s_1 a_1 + s_2 a_2 + \ldots + s_n a_n}{S} \tag{10-15}$$

where,

$s_{1,2,\ldots n}$ are the individual boundary surface areas in ft^2,

$\bar{a}_{1,\,2,\ldots,\,n}$ are the individual absorption coefficients of the individual boundary surface areas,

S is the total boundary surface area in ft^2.

The reflected energy is $1 - \bar{a}$.

Table 10-4 gives typical absorption coefficients for common materials. These coefficients are used to calculate the absorption of boundary surfaces (walls, floors, ceilings, etc).

Table 10-5 gives typical absorption units in sabins rather than percentage figures. Sabins are either in per-unit figures or in units per length.

Finally, the room will possess a reverberation time, RT_{60}. This is the time in seconds that it will take a steady-state sound, once its input power is terminated, to attenuate 60 dB. For the sake of illustration, assume a room with the following characteristics:

$V = 500{,}000\ \text{ft}^3$,

$S = 42{,}500\ \text{ft}^2$,

Table 10-4. Sound Absorption Coefficients of General Building Materials and Furnishings

Materials	Coefficients					
	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz
Acoustical plaster (“Zonolite”)						
½ in. thick trowel application	0.31	0.32	0.52	0.81	0.88	0.84
1 in. thick trowel application	0.25	0.45	0.78	0.92	0.89	0.87
Acoustile, surface glazed and perforated structural clay tile, perforate surface backed with 4 in. glass fiber blanket of 1 lb/ft^2 density	0.26	0.57	0.63	0.96	0.44	0.56
Air (Sabins per 1000 ft$^{3)}$					2.3	7.2
Brick, unglazed	0.03	0.03	0.03	0.04	0.05	0.07
Brick, unglazed, painted	0.01	0.01	0.02	0.02	0.02	0.03
Carpet, heavy						
on concrete	0.02	0.06	0.14	0.37	0.60	0.65
on 40 oz hairfelt or foam rubber with impermeable latex backing	0.08	0.24	0.57	0.69	0.71	0.73
on 40 oz hairfelt or foam rubber						
40 oz hairfelt or foam rubber	0.08	0.27	0.39	0.34	0.48	0.63
Concrete block						
coarse	0.36	0.44	0.31	0.29	0.39	0.25
painted	0.10	0.05	0.06	0.07	0.09	0.08
Fabrics						
light velour, 10 oz/yd^2, hung straight in contact with wall	0.03	0.04	0.11	0.17	0.24	0.35
medium velour, 10 oz/yd^2, draped to half area	0.07	0.31	0.49	0.75	0.70	0.60
heavy velour, 18 oz/s yd^2 draped to half area	0.14	0.35	0.55	0.72	0.70	0.65
Fiberboards, ½ in. normal soft, mounted against solid backing						
unpainted	0.05	0.10	0.15	0.25	0.30	0.3
some painted	0.05	0.10	0.10	0.10	0.10	0.15
Fiberboards, ½ in. normal soft, mounted over 1 in. air space						
unpainted	0.30		0.15		0.10	
some painted	0.30		0.15		0.10	
Fiberglass insulation blankets						
AF100, 1 in., mounting # 4	0.07	0.23	0.42	0.77	0.73	0.70
AF100, 2 in., mounting # 4	0.19	0.51	0.79	0.92	0.82	0.78
AF530, 1 in., mounting # 4	0.09	0.25	0.60	0.81	0.75	0.74
AF530, 2 in., mounting # 4	0.20	0.56	0.89	0.93	0.84	0.80
AF530, 4 in., mounting # 4	0.39	0.91	0.99	0.98	0.93	0.88
Flexboard, 3/16 in. unperforated cement asbestos board mounted over 2 in. air space	0.18	0.11	0.09	0.07	0.03	0.03
Floors						
concrete or terrazzo	0.01	0.01	0.015	0.02	0.02	0.02
linoleum, asphalt, rubber, or cork tile on concrete	0.02	0.03	0.03	0.03	0.03	0.02
wood	0.15	0.11	0.10	0.07	0.06	0.07
wood parquet in asphalt on concrete	0.04	0.04	0.07	0.06	0.06	0.07
Geoacoustic, 13½ in. × 13½ in., 2 in. thick cellular glass tile installed 32 in. o.c. per unit	0.13	0.74	2.35	2.53	2.03	1.73
Glass						
large panes of heavy plate glass	0.18	0.06	0.04	0.03	0.02	0.02
ordinary window glass	0.35	0.25	0.18	0.12	0.07	0.04
Gypsum board, ½ in. nailed to 2 in. × 4 in., 16 in. o.c.	0.29	0.10	0.05	0.04	0.07	0.09
Hardboard panel, ⅛ in., 1 lb/ft^2 with bituminous roofing felt stuck to back, mounted over 2 in. air space	0.90	0.45	0.25	0.15	0.10	0.10
Marble or glazed tile	0.01	0.01	0.01	0.01	0.02	0.02
Masonite, ½ in., mounted over 1 in. air space	0.12	0.28	0.19	0.18	0.19	0.15

Table 10-4. (cont.) Sound Absorption Coefficients of General Building Materials and Furnishings

Materials	Coefficients					
	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz
Mineral or glass wool blanket, 1 in., 5-15 lb/ft^2 density mounted against solid backing						
covered with open weave fabric	0.15	0.35	0.70	0.85	0.90	0.90
covered with 5% perforated hardboard	0.10	0.35	0.85	0.85	0.35	0.15
covered with 10% perforated or 20% slotted hardboard	0.15	0.30	0.75	0.85	0.75	0.40
Mineral or glass wool blanket, 2 in., 5-15 lb/ft^2 density mounted over 1 in. air space						
covered with open weave fabric	0.35	0.70	0.90	0.90	0.95	0.90
covered with 10% perforated or 20% slotted hardboard	0.40	0.80	0.90	0.85	0.75	
Openings						
stage, depending on furnishings				0.25–0.75		
deep balcony, upholstered seats				0.50–1.00		
grills, ventilating				0.15–0.50		
Plaster, gypsum or lime						
smooth finish, on tile or brick	0.013	0.015	0.02	0.03	0.04	0.05
rough finish on lath	0.02	0.03	0.04	0.05	0.04	0.03
smooth finish on lath	0.02	0.02	0.03	0.04	0.04	0.03
Plywood panels						
2 in., glued to 2½ in. thick plaster wall on metal lath	0.05		0.05		0.02	
¼ in., mounted over 3 in. air space, with 1 in. glassfiber batts right behind the panel	0.60	0.30	0.10	0.09	0.09	0.09
⅜ in.	0.28	0.22	0.17	0.09	0.10	0.11
Rockwool blanket, 2 in. thick batt (Semi-Thik)						
mounted against solid backing	0.34	0.52	0.94	0.83	0.81	0.69
mounted over 1 in. air space	0.36	0.62	0.99	0.92	0.92	0.86
mounted over 2 in. air space	0.31	0.70	0.99	0.98	0.92	0.84
Rockwool blanket, 2 in. thick batt (Semi-Thik),covered with 3⁄16 in. thick perforated cement-asbestos board (Transite), 11% open area						
mounted against solid backing	0.23	0.53	0.99	0.91	0.62	0.84
mounted over 1 in. air space	0.39	0.77	0.99	0.83	0.58	0.50
mounted over 2 in. air space	0.39	0.67	0.99	0.92	0.58	0.48
Rockwall blanket, 4 in. thick batt (Full-Thik)						
mounted against solid backing	0.28	0.59	0.88	0.88	0.88	0.72
mounted over 1 in. air space	0.41	0.81	0.99	0.99	0.92	0.83
mounted over 2 in. air space	0.52	0.89	0.99	0.98	0.94	0.86
Rockwool blanket, 4 in. thick batt (Full-Thik), covered with 3⁄16 in. thick perforated cement-asbestos board (Transite), 11% open area						
mounted against solid backing	0.50	0.88	0.99	0.75	0.56	0.45
mounted over 1 in. air space	0.44	0.88	0.99	0.88	0.70	0.30
mounted over 2 in. air space	0.62	0.89	0.99	0.92	0.70	0.58
Roofing felt, bituminous, two layers, 0.8 lb/ft^2, mounted over 10 in. air space	0.50	0.30	0.20	0.10	0.10	0.10
Spincoustic blanket						
1 in., mounted against solid backing	0.13	0.38	0.79	0.92	0.83	0.76
2 in., mounted against solid backing	0.45	0.77	0.99	0.99	0.91	0.78
Spincoustic blanket, 2 in., covered with 3⁄16 in. perforated cement-asbestos board (Transite), 11% open area	0.25	0.80	0.99	0.93	0.72	0.58
Sprayed "Limpet" asbestos						
¾ in., 1 coat, unpainted on solid backing	0.08	0.19	0.70	0.89	0.95	0.85
1 in., 1 coat, unpainted on solid backing	0.30	0.42	0.74	0.96	0.95	0.96

Table 10-4. (cont.) Sound Absorption Coefficients of General Building Materials and Furnishings

Materials	Coefficients 125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz
¾ in., 1 coat, unpainted on metal lath	0.41	0.88	0.90	0.88	0.91	0.81
Transite, 3⁄16 in. perforated, cement-asbestos board, 11% open area						
mounted against solid backing	0.01	0.02	0.02	0.05	0.03	0.08
mounted over 1 in. air space	0.02	0.05	0.06	0.16	0.19	0.12
mounted over 2 in. air space	0.02	0.03	0.12	0.27	0.06	0.09
mounted over 4 in. air space	0.02	0.05	0.17	0.17	0.11	0.17
paper-backed board, mounted over 4 in. air space	0.34	0.57	0.77	0.79	0.43	0.45
Water surface, as in a swimming pool	0.008	0.008	0.013	0.015	0.02	0.025
Wood paneling, ⅜ in. to ½ in. thick, mounted over 2 in. to 4 in. air space	0.30	0.25	0.20	0.17	0.15	0.10

Table 10-5. Absorption of Seats and Audience*

Materials	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz
Audience, seated, depending on spacing and upholstery of seats	2.5–4.0	3.5–5.0	4.0–5.5	4.5–6.5	5.0–7.0	4.5–7.0
Seats						
heavily upholstered with fabric	1.5–3.5	3.5–4.5	4.0–5.0	4.0–5.5	3.5–5.5	3.5–4.5
heavily upholstered with leather, plastic, etc.	2.5–3.5	3.0–4.5	3.0–4.0	2.0–4.0	1.5–4.0	1.0–3.0
lightly upholstered with leather, plastic, etc.			1.5–2.0			
wood veneer, no upholstery	0.15	0.20	0.25	0.30	0.50	0.50
Wood pews						
no cushions, per 18 in. length			0.40			
cushioned, per 18 in. length			1.8–2.3			

*Values given are in sabins per person or unit of seating

$\bar{a} = 0.128.$

therefore the RT_{60} is

$$RT_{60} = \frac{0.049V}{Sa}$$
$$= 4.5 \text{ s}$$

(See Chapter 12 *Large Room Acoustics*, for a more detailed development of this subject.)

10.16.1 The Mean Free Path (MFP)

The mean free path is the average distance between reflections in a space. For our sample space:

$$MFP = 4\frac{V}{S}$$
$$= 4\left(\frac{500{,}000}{42{,}500}\right)$$
$$= 47 \text{ ft}$$

If a sound is generated in the sample space, part of it will travel directly to a listener and undergo inverse-square-law level change on its way. Some more of it will arrive after having traveled first to some reflecting surface, and still more will finally arrive having undergone several successive reflections (each 47 ft apart on the average). Each of these signals will have had more attenuation at some frequencies than at others due to divergence, absorption, reflection, refraction, diffraction, etc.

We can look at this situation in a different manner. Each sound made will have traveled 4.5 s × 1130 ft/s, or 5085 ft. Since the mean free path is 47 ft, then we can assume each sound underwent approximately 108 reflections in this sample space before becoming inaudible. The result is a lot different than hearing the sound just once.

10.16.2 The Build-Up of the Reverberant Sound Field

Fig. 10-13 shows the paths of direct sound and several reflected sound waves in a concert hall.

Reflections also occur from balcony faces, rear wall, niches, and any other reflecting surfaces. We can obtain a chart such as that shown in Fig. 10-14 if we plot the amplitude of a short-duration signal vertically and the time interval horizontally. This diagram shows that at listener's ears, the sound that travels directly from the performer arrives first, and after a gap, reflections from the walls, ceiling, stage enclosure, and other reflecting surfaces arrive in rapid succession. The height of a bar suggests the loudness of the sound. This kind of diagram is called a reflection pattern. The initial-signal-delay gap (*ISDG*) can be measured from it.

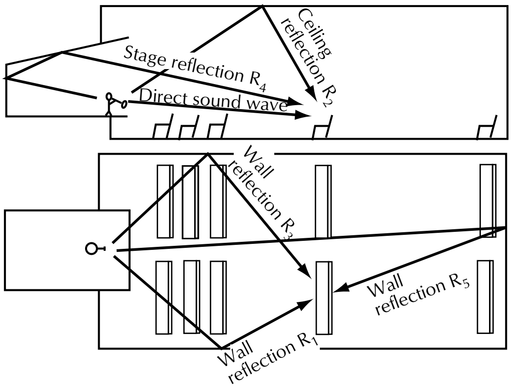


Figure 10-13. Sound paths in a concert hall.

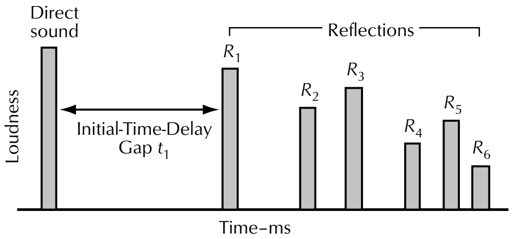
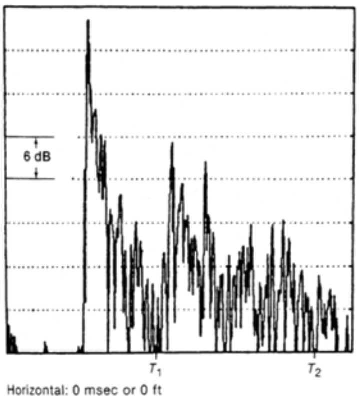


Figure 10-14. Time relationship of direct and reflected sounds.

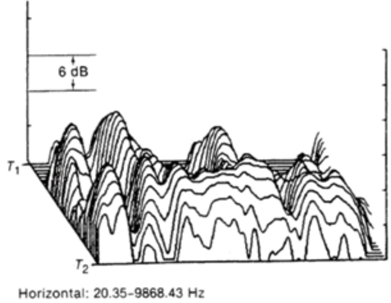
Fig. 10-14 illustrates the decay of the reverberant field. Here the direct sound enters at the left of the diagram. The initial-signal-delay gap is followed by a succession of sound reflections. The reverberation time of the room is defined as the length of time required for the reverberant sound to decay 60 dB.

We will encounter the effects of delay versus attenuation again when we approach the calculation of articulation losses of consonants in speech.

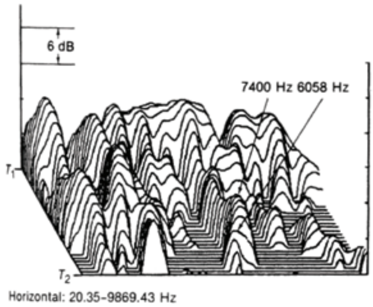
Fig. 10-15 shows measurements from an analyzer made in both large and small rooms. Fig. 10-16 shows that the sound arriving at the listener has at least three distinct divisions:



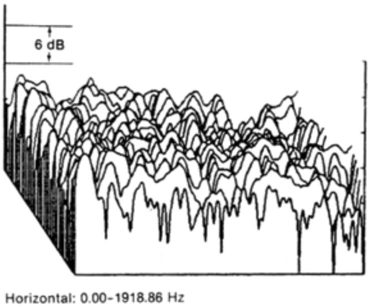
A. Envelope Time Curve (ETC) of a small room showing lack of a dense field of reflections



B. Small room without reverberant sound field but with room modes



C. Small room without reverberant sound field showing decay side of room modes



D. Large room with reverberant sound field

Figure 10-15. Vivid proof that there is a fundamental difference between a small reverberant space and a large reverberant hall.

1. The direct sound level L_D.
2. The early reflections level L_{RE}.
3. The reverberant sound level L_R.

The direct sound, by definition, undergoes no reflections and follows inverse-square-law level change. The reverberant sound tends to remain at a constant level if the sound source continues to put energy into the room at a reasonably regular rate. This gives rise to a number of basic sound fields, Fig. 10-17:

1. The near field.
2. The far free field.
3. The far reverberant field.

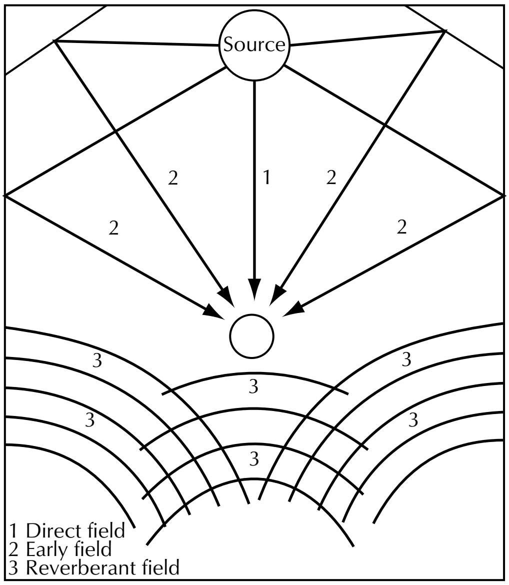


Figure 10-16. Comparison of direct, early, and reverberant sound fields in an auditorium (reflection angles adjusted for purposes of illustration).

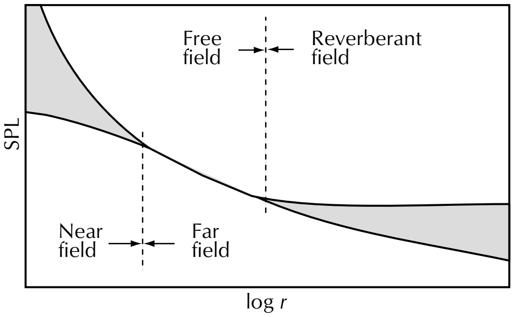


Figure 10-17. Graphic representation of near field, free field, and reverberant field.

The near field does not behave predictably in terms of L_P versus distance because the particle velocity is not necessarily in the direction of travel of the wave, and an appreciable tangential velocity component may exist at any point. This is why measurements are usually not made closer than twice the largest dimension of the sound source. In the far free field, inverse-square-law level change prevails. In the far reverberant field, or diffuse field, the sound-energy density is very nearly uniform. Measuring low-frequency loudspeakers is an exception to the rule, and such measurements are often made in the pressure response zone of the device.

10.17 Conclusion

The study of acoustics for sound system engineers divides into outdoors and indoors with indoor acoustics again divided into large room acoustics and small room acoustics. Classical Sabinian acoustics are rapidly being refined where applicable, discarded where misapplied, and reexamined where the "fine structure of reverberation" is the meaningful parameter. The digital computer has fueled basic research into the mathematics of enclosed spaces and modern analyzers have served to verify or deny the validity of the theories put forward.

Bibliography

Acoustical Materials Assoc. *The Use of Architectural Materials—Theory and Practice,* 335 East 45 ST., NewYork, N.Y. 10017.

D. Davis and C. Davis. "What Reverberation Is and What It Is Not," *Syn-Aud-Con Tech Topic*, Vol. 12, No. 13 (1985).

L. E. Kinsler and A. R. Frey. *Fundamentals of Acoustics*, 2nd ed. New York: Wiley, 1962.

V. O. Knudsen and C. M. Harris. *Acoustical Designing in Architecture*, New York: Wiley, 1950

H. Kuttruff. *Room Acoustics*. New York: Halstead Press, 1973.

B. R. Lindsay. *Acoustics—Historical and Philosophical Development*. Stroudsburg, Penn.: Dowden, Hutchinson & Ross, 1973.

R. MacKenzie, Ed. *Auditorium Acoustics*. London: Applied Science Publishers, 1975.

H. F. Olson. *Music, Physics, and Engineering*. New York: Dover, 1966.

A. D. Pierce. *Acoustics: An Introduction to Its Physical Principles and Applications*. New York: McGraw-Hill, 1981.

J. R. Pierce. *The Science of Musical Sound*. New York: Scientific American Books, 1983.

T. D. Rossing. *The Science of Sound*. Reading, Mass.: Addison-Wesley, 1982.

P. E. Sabine. *Acoustics and Architecture*. New York: McGraw-Hill, 1932.

W. C. Sabine. *Collected Papers on Acoustics*. Cambridge, Mass.: Harvard Univ. Press, 1922.

L. J. Sivian, H. K. Dunn, and S. D. White. “Absolute Amplitudes and Spectra of Certain Musical Instruments and Orchestras,” *IRE Trans. on Audio* (May-June 1959), pp. 47-75.

J. W. Strutt and B. Rayleigh. *The Theory of Sound*, Vols. I and 11, 2nd ed. New York: Dover, 1945.

Chapter 11

Audio and Acoustic Measurements

by Don Davis

11.1 Acoustic Analysis Sans Instrumentation . . . 191
11.2 Initial Parameters . . . 191
11.3 Acoustic Tests of Sound Systems . . . 192
  Acoustic Test Signals . . . 192
  Where to Place the Microphone . . . 193
  Measurement Analyzers . . . 193
11.4 Examining AC Outlets . . . 193
  First, Look and Listen . . . 195
11.5 The ETC Plot . . . 195
  Impulse Response . . . 195
  The Heyser Spiral . . . 196
  The Magnitude and Phase Response . . . 196
  Three Parameter Measurements . . . 198
  The Nyquist Plot . . . 200
  The Polar Envelope Time (PET) Plots . . . 201
  History of Polar Time Measurements . . . 201
11.6 Site Surveys and Noise Criteria Curves . . . 203
  Constant Percentage Bandwidth Analysis . . . 203
11.7 An Improper Use of Real Time Analysis . . . 203
11.8 Evaluation of Listener Response . . . 205
11.9 Fractional Bandwidth Filter Analyzers . . . 206
  Series . . . 207
  Other Uses . . . 207
  Useful Tools . . . 208
  Using the Decade Exponents . . . 208
  Decibels . . . 208
  Label Frequencies . . . 208
11.10 Measuring Electromagnetic Pollution . . . 209
11.11 Conclusion . . . 211
Bibliography . . . 211

In order to better understand what we hear we often turn to measurements. As one authority in theory once remarked, “When the number of variables approached an order of magnitude, I turn in despair to my measurement apparatus.”

The finest acoustical measurement apparatus available cannot duplicate what a trained human listener can achieve. If we examine an unknown signal with all extant equipment we can’t tell if its music, noise, speech, or gibberish, but a $2 loudspeaker allows the trained human listener to tell which, and if speech, what language.

Instrumentation is used to measure room parameters before the design begins, to compute design factors, to install the system, and, finally, to operate and maintain the system. The greatest single division between professional work and nonprofessional work in the system business is the use and understanding of basic audio and acoustic instrumentation.

The following quote is pertinent to the intent of this chapter:

> *I often say that when you can measure what you are speaking about, and can express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind. It may be the beginning of knowledge but you have scarcely advanced to the stage of science whatever the matter may be...*

Lord Kelvin 1824-1907

11.1 Acoustic Analysis Sans Instrumentation

What defines a difficult acoustic space? Excessive noise, too-long a reverberation time at the wrong frequencies, faulty geometry resulting in focused high-level reflections, inappropriate materials for the function at hand such as marble in conference rooms, inappropriate physical locations such as recording studios too near railroads and airports.

These are but a few of the problems that owners, architects, and enthusiastic volunteer consultants can generate.

When an acoustic space physically exists the number one priority is to visit it, especially when it is in its normal use, and listen to it using the two channel analyzer our maker provided us, namely two ears, a rotating head, and a right and left hemisphere in our brain. Cupping our earlobes provides us with a highly directional antenna that can by walking through anomalies in the room locate areas for more detailed analysis by conventional equipment. Carrying a small square of an absorptive material, such as Sonex, allows focused reflections to be blocked for acoustic evaluation.

Walking large arenas will often result in the detection of localized, very high level noise sources from air-conditioning ducts, inadequate isolation of machinery spaces, flanking paths from other spaces, all of which might not interfere with the main audience, but guarantees violent criticism from the customers seated in those areas.

A careful walk-through a church auditorium often reveals seats not only acoustically isolated, but visually isolated, such as under balconies or in various side areas. It is disturbing to see designers that are totally unaware of the differences in church liturgy, and proceed to design to their own church’s standard, regardless of the inappropriateness to the church at hand.

In one case when I entered the church the organist told me that it was the deadest space he had ever played in, followed a few minutes later by the minister saying it was the most reverberant church he had ever preached in. Walkabout analysis revealed that the bass frequencies were being sucked out of the room by the beautifully paneled ceiling into the attic space, and that the hard oak pews were making it difficult for speech when a full audience wasn't present. After detailed conventional analysis, stiffening the ceiling panels, and providing pew cushions, resulted in a completely satisfactory space for both the organist and the minister.

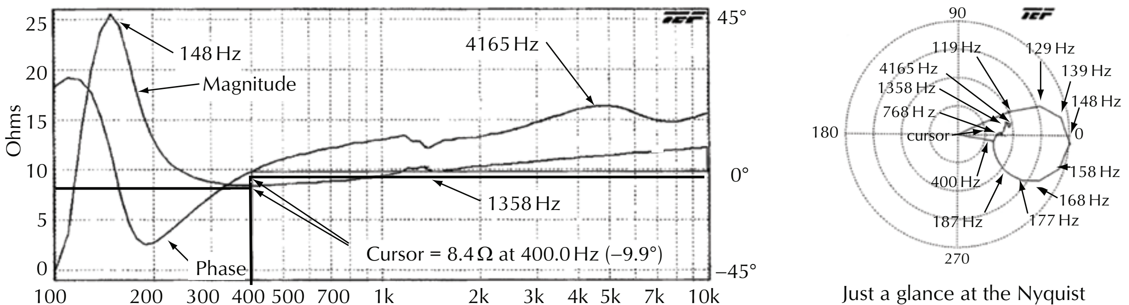
In one case when asked what they could do to remedy an extremely difficult space a famous consultant replied “pray for a high wind” as the space was in tornado country. Fortunately in many cases what disturbs highly trained listeners will, with minor modifications, satisfy the ultimate users. That is reason enough to proceed step-by-step in the correction of existing difficult spaces to allow for acceptability to be detected by the ultimate end-user.

Learn and understand what the space is used for, listen to it in actual use, and whenever possible, in the company of one of the ultimate users so that you can see his or her response to the anomalies that you detect, and then proceed with confidence to more detailed measurements that your ultimate user will then be more likely to understand and accept.

11.2 Initial Parameters

To make a measurement, choose the initial parameters by one of the three following techniques:

1. Experience with similar devices.



Just a glance at the Nyquist display reveals which frequencies are predominately resistive, capacitance, or inductive impedance values.

$Z = |Z|\, e^{j\theta}$
where,
$|Z|$ is the magnitude of the impedance
θ is the angle of the impedance

$\text{Reactance} = |Z| \sin\theta$
$\text{Resistance} = |Z| \cos\theta$
$\text{Angle} = \arctan\left(\frac{\text{Reactance}}{\text{Resistance}}\right)$

Frequency	400 Hz
Phase angle	–9.88°
Impedance	8.41 Ω
Resistance	8.29 Ω
Reactance	–1.44 Ω

Figure 11-1. Loudspeaker impedances.

2. Mathematical analysis of the device and its most likely performance.
3. Cut and try experimentation.

Component designers often, justifiably, use Step 3. System designers must not. System designers need to specify proven, trusted components. Systems are complex by their very nature of combining components from many manufacturers. To increase that complexity with untried components is irresponsible.

11.3 Acoustic Tests of Sound Systems

Once all the electrical tests of the sound system are completed and any electrical problems are corrected meaningful acoustic tests can be performed to verify:

1. Output levels and areas of coverage for individual transducers comprising the acoustic output of the system.
2. The phase and polarity of the individual transducers.
3. The signal synchronization between different transducers sharing identical areas of coverage, i.e., overlap zones.
4. The absence of any undesired spurious energy returns from any reflective surface.
5. The measurement of the relationship of $L_D - L_R$ to confirm the $\%Al_{CONS}$ at selected audience locations.
6. The equalization of L_D as required.
7. The loudspeaker impedances, see Fig. 11-1.

11.3.1 Acoustic Test Signals

Sound engineers have available many different test sources:

Music and speech. Excellent if the listener is highly trained—a rarity.

The steady-state sine wave. While perhaps our most useful electrical signal, it is rarely useful in acoustic tests.

Swept sine wave. This source is our single most useful test signal.

Random noise. White, pink, USASI (or ANSI), and other special forms of noise are useful for magnitude measurements, but they pose too great a complexity in the acquisition of phase measurements.

Impulse sources. These sources represent the worst possible choice for acoustic measurements, especially when used in conjunction with FFT analysis. They offer the least effective use possible of the test signal's energy.

The starting point for any serious acoustic measurement system is the calibrated measurement microphone. Fig. 11-2 allows the comparison of the human ear with a quality electret microphone. Notice that "intelligent design" has allowed the microphone, in some parameters, to exceed the human ear and in others to equal it for all practical purposes. One of the authors worked for a company that built a microphone capable of $L_P = 220\,\text{dB}$ (for measuring the pressure wave from a hydrogen bomb explosion which produced a spike well above atmospheric pressure).

Specification	*Ear*	*Mic**	*Units*
Size	12	0.17	cm^3
Power Consumption	50	25	µW
Vibration Sensitivity (1 g)	100	75	dB SPL
Shock Resistance	100	20,000	g
Noise Level (A-weighted)	20	20	dB SPL
Overload Level (10% THD)	100	140	dB SPL
Dynamic Range (Pure Tone)	100	140	dB
Acoustic Input Z (low frequency)	1.4	0.03	cm^3
Frequency response	25–16,000	10–25,000	Hz

---

*Small electret

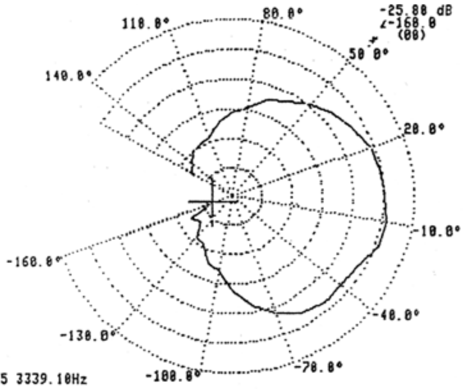
Figure 11-2. Comparison of an ear and a microphone. (*Courtesy Dr. Mead Killion.*)

11.3.2 Where to Place the Microphone

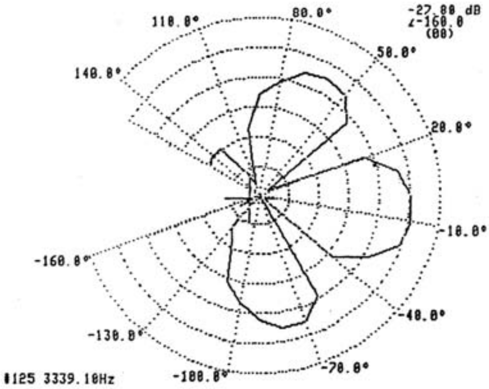
This frequent query has an easy answer: where your ears tell you there is a problem, or where you need to look at the radiation pattern for adjustment. For example, the point equally distant from two loudspeakers where you desire to "signal align" them for minimum polar response interferences, see Figs. 11-3A and 11-3B. Another example, the area where your ears tell you that intelligibility has suffered for no visually apparent reason and the measurement of the Envelope Time Curve, ETC, reveals an unexpected focused reflection damaging the direct sound level in that area.

11.3.3 Measurement Analyzers

The authors prefer Richard C. Heyser's analysis system as exemplified in a TEF instrument. This is not to deprecate other devices but is the result of the superior signal-to-noise parameters so vital to field measurements. The Heyser Integral Transform is unique, Fig. 11-4. While it has yet to realize its full potential in real instruments, its embodiment in what's currently available has radically changed how we measure. Indeed the frequency modulation function identified by Heyser has found an embodiment in HP's modulation domain analysis, Fig. 11-5A, 11-5B, and 11-5C.

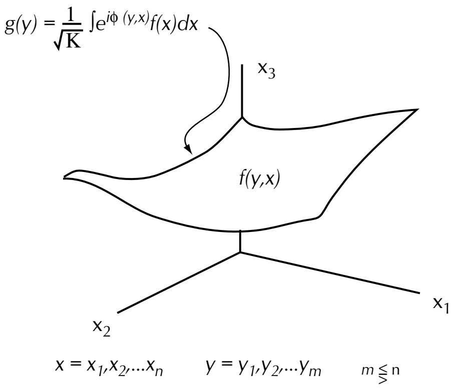


A. Horizontal polar response of two loudspeakers one stacked on top of the other and in physical alignment.



B. Horizontal polar response of two loudspeakers one stacked on top of the other and out of physical alignment by 3 inches.

Figure 11-3. The effect of mis-alignment on loudspeaker output.

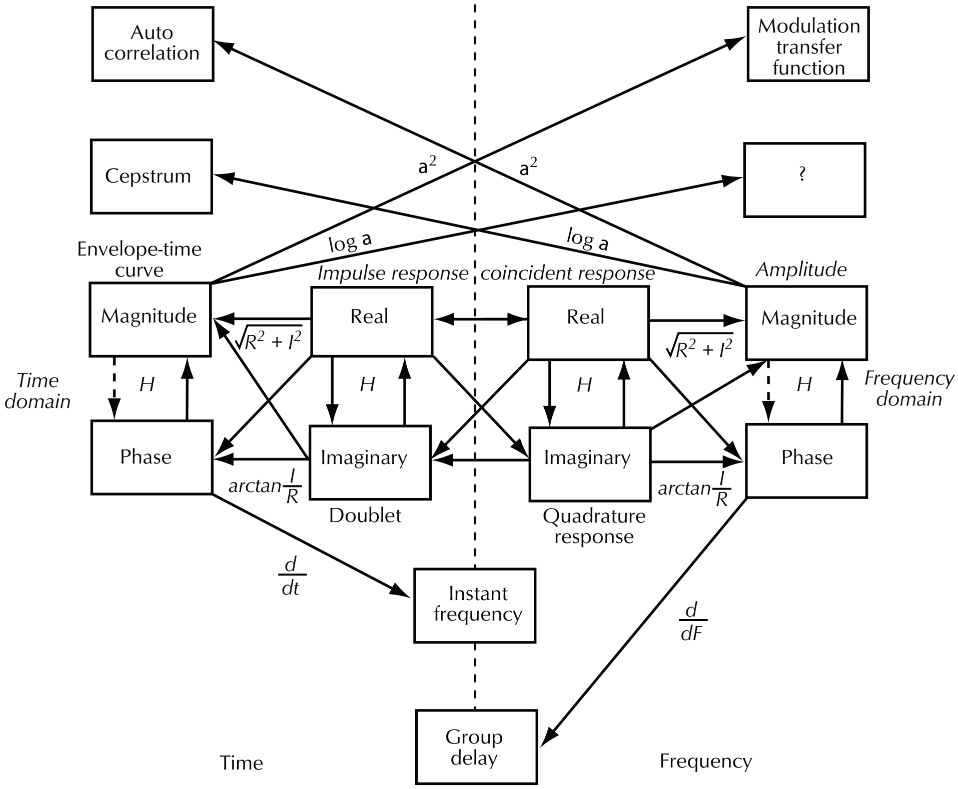

$g(y) = \frac{1}{\sqrt{K}} \int e^{i\phi(y,x)} f(x) dx$
x_3
$f(y,x)$
x_2
x_1
$x = x_1, x_2, \ldots x_n$
$y = y_1, y_2, \ldots y_m$
$m \leq n$

f(*y*,*x*) is a hypersurface, defined in parameters *y* and expressed over all *x*.

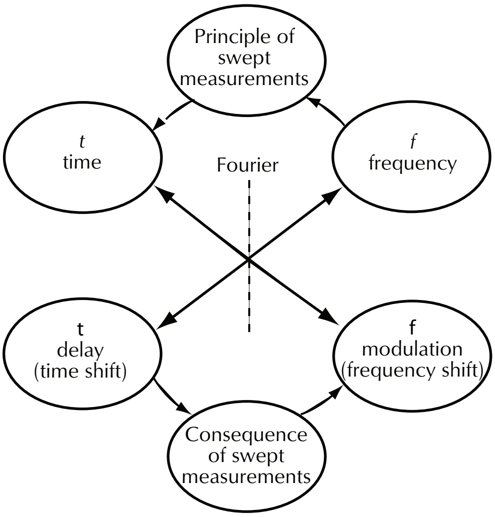
Figure 11-4. Heyser Integral Transform.

11.4 Examining AC Outlets

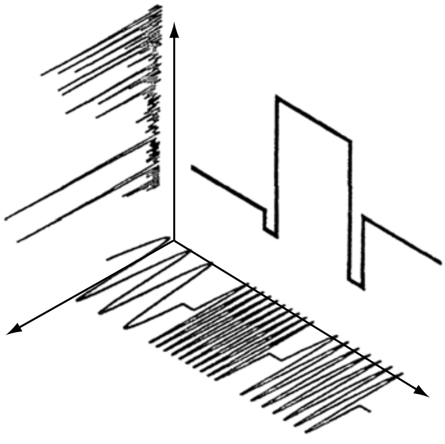
Prior to plugging any valuable equipment into an ac outlet it should be examined for unsuspected dc, (in one case a light dimming circuit,in a motel meeting room, was also wired to a three wire wall receptacle), correct ac voltage, proper frequency, hot, neutral and ground conductor impedances and proper wiring in a three wire receptacles. GFCIs (ground fault circuit interrupters) and AFCIs (arc fault circuit interrupters) should have their proper performance tested and verified. When anomalies appear in such


Auto correlation
Modulation transfer function
Cepstrum
?
a^2
a^2
log a
log a
Envelope-time curve
Impulse response
coincident response
Amplitude
Magnitude
Real
Real
Magnitude
$\sqrt{R^2 + I^2}$
$\sqrt{R^2 + I^2}$
Time domain
H
H
H
H
Frequency domain
Phase
Imaginary
Imaginary
Phase
$\arctan\frac{I}{R}$
$\arctan\frac{I}{R}$
Doublet
Quadrature response
$\frac{d}{dt}$
Instant frequency
$\frac{d}{dF}$
Group delay
Time
Frequency

A. TEF measurements.


Principle of swept measurements
t
time
Fourier
f
frequency
t
delay
(time shift)
f
modulation
(frequency shift)
Consequence of swept measurements

B. When the frequency response of a time invariant system is measured as a function of time the (time) delay response is converted to a (frequency) modulation function.



C. The same signal can be represented in the time domain on an oscilloscope (bottom) and in the frequency domain on a spectrum analyzer (left). Hewlett-Packard's 5371A frequency and time-interval analyzer shows the signal's frequency against time, inaugurating what HP calls the modulation domain. The analyzer simplifies such measurements as timing jitter, frequency drift, and modulation on communications signals.

Figure 11-5. Processing TEF signals.

testing a careful examination of the wiring system should be undertaken both for safety reasons and for control of electrical noise in your system.

I can recall one case where a student of mine, a sound contractor, called for the wiring inspection of his new building. When the inspector arrived he was handed a check for his work and turned to go out the door. The student asked, “Aren’t you going to check anything?” The inspector gave the check a long look and said, “It looks all right to me.”

11.4.1 First, Look and Listen

Upon arrival at a measurement site, first look and listen. We should visually inspect the site and then listen, sans sound system. Take time to walk the audience area and listen to a live talker standing where the performer will be. Listen for noise masking of the talker—echoes, focused reflections, and strange level dropouts, i.e., cancellations. Cupping your ears allows some directional discrimination with regard to reflections. This exercise helps to quickly ascertain if it's the room or the system or both that need correction. Many existing sound systems in difficulty exist in rooms where the unaided voice can be heard clearly.

A first look and listen allows identification of logical measurement points. Once a given point is selected make a global Envelope Time Curve, ETC. The time base needs to be at least twice the room's largest dimension, or longer than the "by the ear" estimate of the room's RT_{60}. The source used should be one loudspeaker from the array that has the microphone in its path. If inspecting the room before the sound installation is possible, use a test loudspeaker suitable for such a space or if that is not possible, at a minimum, use one with a known Q. Whenever possible, mount the loudspeaker, via a portable hoist, in a logical location for a proposed system.

A global view of the time domain insures that late arriving energy is not overlooked and at the same time allows estimating the shorter time scales that will be employed to obtain more detail. A rule to remember is that you can truncate long time to shorter time but not the reverse. From the first global measurement, we looked at the Heyser Spiral. See Fig. 11-6A, followed by the ETC in Fig. 11-6B. See Fig. 11-6C for the truncated ETC of a shorter interval which revealed a missynchronized package loudspeaker that was non-minimum phase. See Fig. 11-6D for the Nyquist, where the curve encircled the origin.

These initial measurements revealed a few of the items needing correction in a minimum amount of time.

11.5 The ETC Plot

The Envelope Time Curve, ETC, is related to a well-established concept in communication theory known as the modulation envelope. The Envelope Time Curve is the magnitude of the analytic signal description of the impulse response.

In the acoustical measurement case, let I represent the impulse response which is a real function of time and let $\overline{I}$ represent the Hilbert Transform of the impulse response. Also, let I_A represent the analytic impulse response. Then,

$$I_A = I + j\overline{I} \tag{11-1}$$

Now, consider the quantity

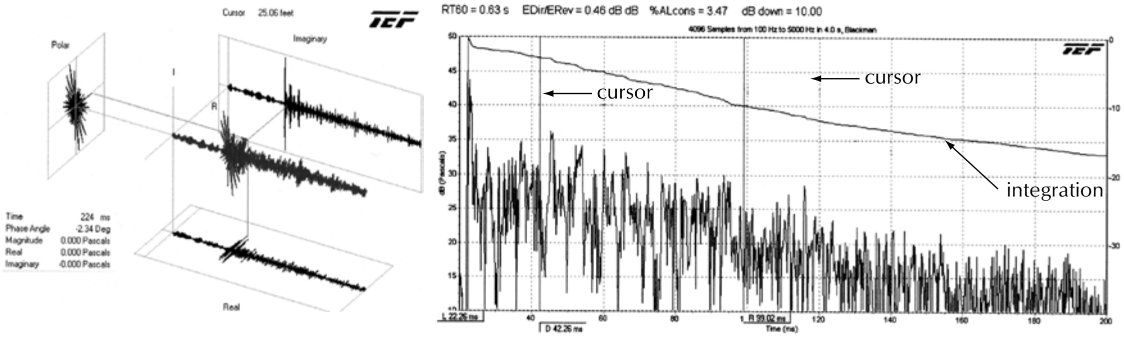
$$20\log\frac{\sqrt{I^2+\overline{I}^2}}{2\times 10^{-5}} = 20\log\frac{|I_A|}{2\times 10^{-5}} \tag{11-2}$$

This is the quantity plotted versus time in forming the curve known as the ETC. This is similar to a smoothed version of the impulse squared response. It has proven its worth over the years in identifying detrimental reflections, locating the desired signal delay corrections for magnitude measurements in the frequency domain, usually "fine tuned" by microsecond adjustments of the phase response, and in examining the density or lack of density of the reflected sound field at any given point in space.

11.5.1 Impulse Response

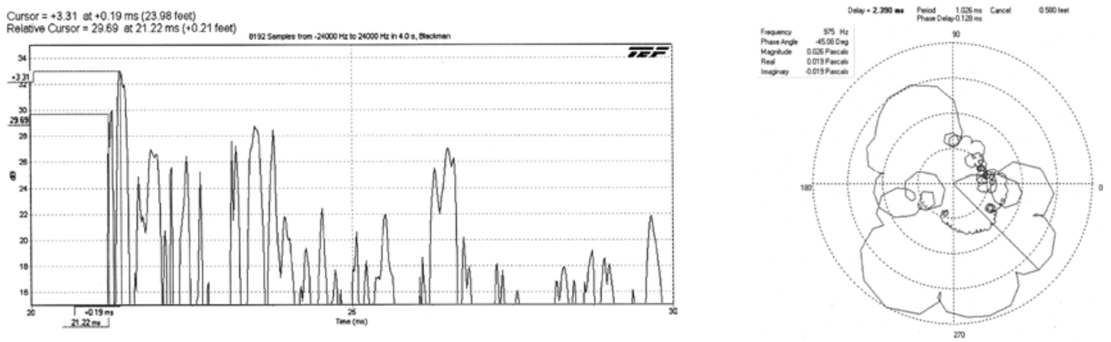
Today impulse responses are acquired in the frequency domain both to address undue stress on the loudspeaker system and to obtain an improved signal-to-noise ratio, *SNR*. It is then inverse Fourier transformed to the time domain where it can be displayed in a number of forms, see Chapter 22 *Signal Processing*, for an explanation of Fourier transform. The Fourier transform takes both the real and the imaginary parts, from the amplitude and phase measurements in the frequency domain, to compute the impulse response in the time domain, Fig. 11-7A. Additionally, the impulse response can be Hilbert transformed to produce the doublet response. The impulse response forms the real part of the complex ETC while the doublet response forms the imaginary part of the complex ETC, Fig. 11-7B. Fig. 11-7C shows the relationship between the real and imaginary parts on the Heyser Spiral.

Modern analyzers provide differing viewpoints of the same information. A good example is the comparison of the log squared amplitude of the impulse response, Fig. 11-7D, to the ETC, Fig. 11-7E. The impulse response contains all the data but the ETC clearly shows some arrivals with greater clarity due to the modulation domain aspect of the envelope.



A. A Heyser spiral of the room in the time domain.

B. The Envelope Time Curve (ETC). Note that the cursors (readout above plot) solve for RT_{60}, $\%AL_{CONS}$, and direct to reflected levels.



C. ETC with shorter time interval than Heyser spiral revealing a non synchronized loudspeaker.

D. Nyquist display showing non minimum phase encirclement of origin.

Figure 11-6. Variety of TEF displays.

11.5.2 The Heyser Spiral

Richard C. Heyser's remarkable insights, so often copied, so seldom acknowledged, that signal acquisition in the frequency domain via a swept sine wave (chirp) tracked by a time offset tracking filter yields vastly superior *SNR* in both the frequency domain and the inverse Fourier transformed time domain.

The easiest visualization of these processes is the Heyser Spiral display. Fig. 11-8A shows the frequency domain Heyser Spiral composed of the complex signal on the frequency axis, the real and imaginary parts shadowed on the appropriate planes, and the Nyquist trace of the complex signal.

An inverse Fourier transform of both the real and imaginary parts in the frequency domain produces the impulse response (real) in the time domain, Fig. 11-8B.

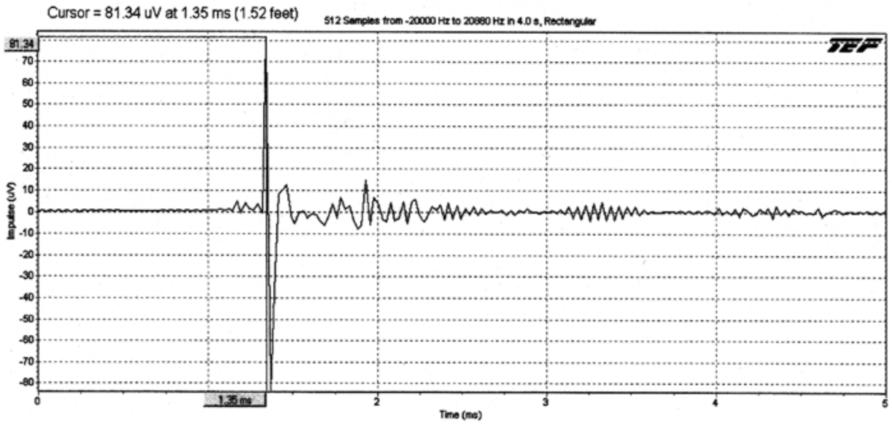
A Hilbert transform of the impulse response (real) produces the doublet response (imaginary), Fig. 11-8C. These real and imaginary parts yield the Envelope Time Curve.

11.5.3 The Magnitude and Phase Response

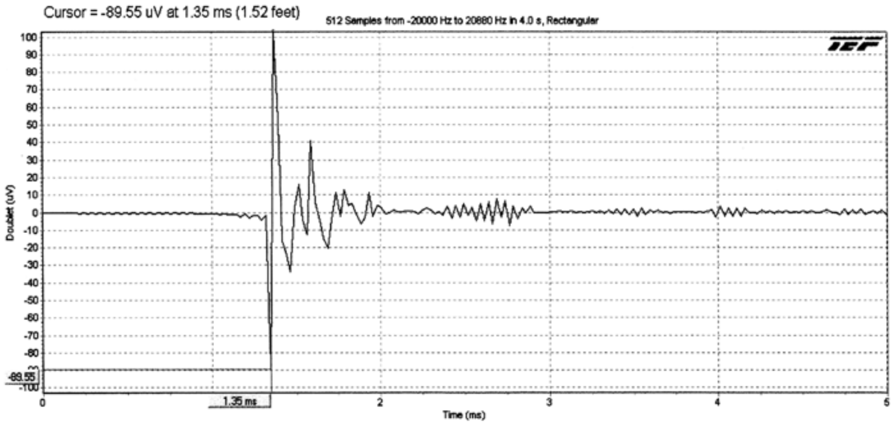
Prior to Heyser, real life data from manufacturers was usually frequency vs. level responses and rarely phase response. The magnitude response is the most familiar measurement to many. It has limited value without the accompanying phase response. The phase response requires "fine tuning" via the microsecond adjustments available in modern analyzers. The adjustment is used to bring the phase response to 0° wherever the magnitude response is uniform. Once this has been done, if the device is minimum phase the peaks and dips on the phase response will be opposite the "slopes" on the magnitude response.

Why all the emphasis on minimum phase response? It is because you cannot apply conventional inverse equalization to the magnitude response unless that portion of the magnitude response is minimum phase. Non-minimum phase usually implies a significant signal delay.

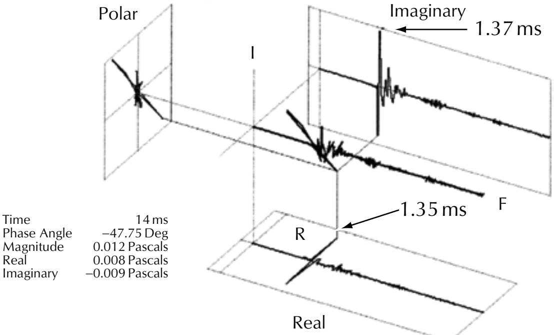
Magnitude response has a vertical decibel scale and a horizontal frequency scale. Phase has a vertical scale in plus or minus degrees and a horizontal frequency scale. In all measurements it is



A. TEF impulse.



B. TEF doublet.

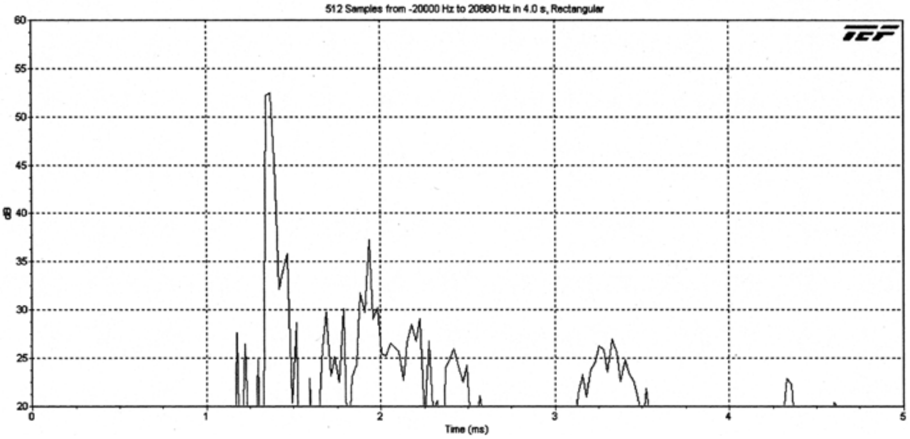


C. Heyser response in the time domain.

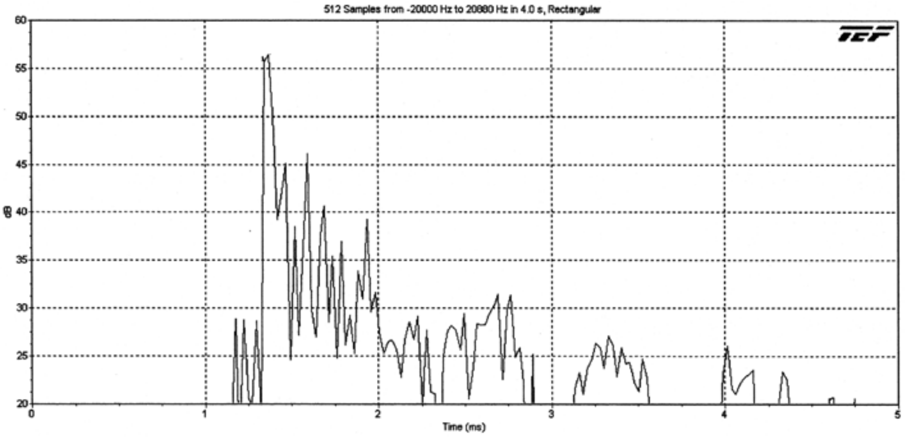
Figure 11-7. Time domain displays.

vital to know what the frequency resolution is. Resolution that is too broad gives optimistic smoothing whereas resolution too narrow includes, in many cases, undesired reflected information.

Measuring phase instead of magnitude provides greater sensitivity and resolution. For example, finding resonant frequencies (phase passes through zero at resonance), the phase response will typically be 10 times more sensitive than the magnitude response. Acoustic delay problems jump out in phase and can be difficult, at best, with magnitude response. The pairing of magnitude and phase, Fig. 11-9 (upper two curves), allows detection of non-minimum phase frequencies—the phase inflec-



D. Log^2 implulse.



E. ETC.

Figure 11-7. (cont) Time domain displays.

tion points don't intersect the center of the magnitude slopes. Further, a flattened phase response over a selected range reveals that the magnitude correction was properly done, Fig. 11-9 (lower two curves).

A non-minimum phase system is one that exhibits an excess delay of the signal over that termed the phase delay. Since an increasing number of audio devices include, either deliberately or accidentally, all-pass components, phase measurements are of ever increasing importance.

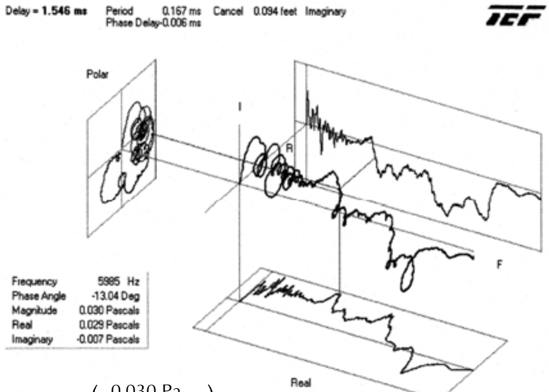
11.5.4 Three Parameter Measurements

In the arsenal of analysis today are the three parameter measurements where the resolution of two parameters is "smeared" to allow a conceptual view of what is occurring. We can choose to compromise the frequency magnitude resolution for higher time resolution, or we can compromise the time resolution for higher frequency magnitude resolution. The typical choice, because we have both ETC and EFC, energy frequency curve, for detailed individual views, is to smear both frequency and time resolutions for a compromise view of what some decaying frequency areas do over time. Typical display parameters are frequency on the horizontal scale, magnitude on the vertical scale, and time on the diagonal scale, Fig. 11-10.

While it always remains true that the reciprocal of the frequency bandwidth determines the time resolution and the reciprocal of the time window determines the frequency resolution, it is possible by "smearing" each parameter to gain an insight into the frequency vs. time behavior of a system, especially when some frequencies are longer in decaying than other frequencies.

$$\Delta f \times \Delta T \geq 1 \qquad (11\text{-}3)$$

where,
Δf is frequency resolution,
ΔT is time resolution.

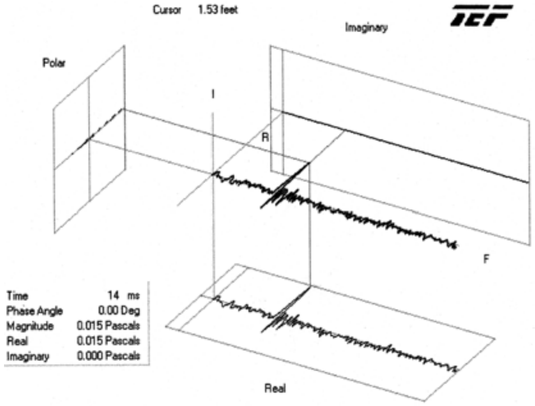


$$20\log\left(\frac{0.030\ \text{Pa}}{0.00002\ \text{Pa}}\right) = 63.5\ \text{dB}$$

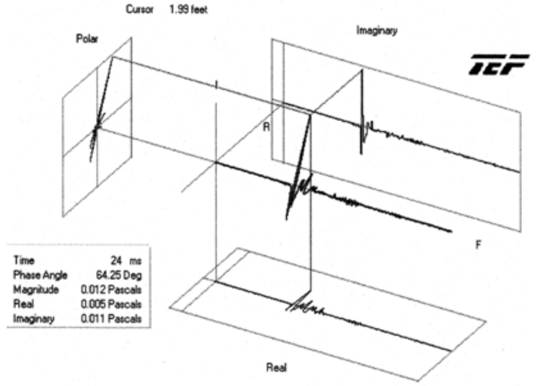
$$\text{atan}\left(\frac{-0.007}{0.029}\right) = -13°$$

$$20\log\left(\frac{\sqrt{(0.029)^2 + (0.007)^2}}{0.00002\ \text{Pa}}\right) = 63.5\ \text{dB}$$

A. Heyser Spiral in the frequency domain.



B. Fourier transform from frequency domain to time domain of real and imaginary parts in the frequency domain yields the impulse response.



$$\text{magnitude} = 20\log\left(\frac{0.012\ \text{Pa}}{0.00002\ \text{Pa}}\right) = 55.6\ \text{dB}$$

$$\text{magnitude in Pa} = \sqrt{(0.029)^2 + (0.007)^2} = 0.030\ \text{Pa}$$

$$\text{phase angle} = \text{atan}\left(\frac{0.011\ \text{Pa}}{0.005\ \text{Pa}}\right) = 65.6\ \text{dB}$$

C. Heyser Spiral in the time domain.

Figure 11-8. Using the Hilbert Transform of a function of time (convolution with 1/πt) yields the imaginary (doublet) in the time domain.

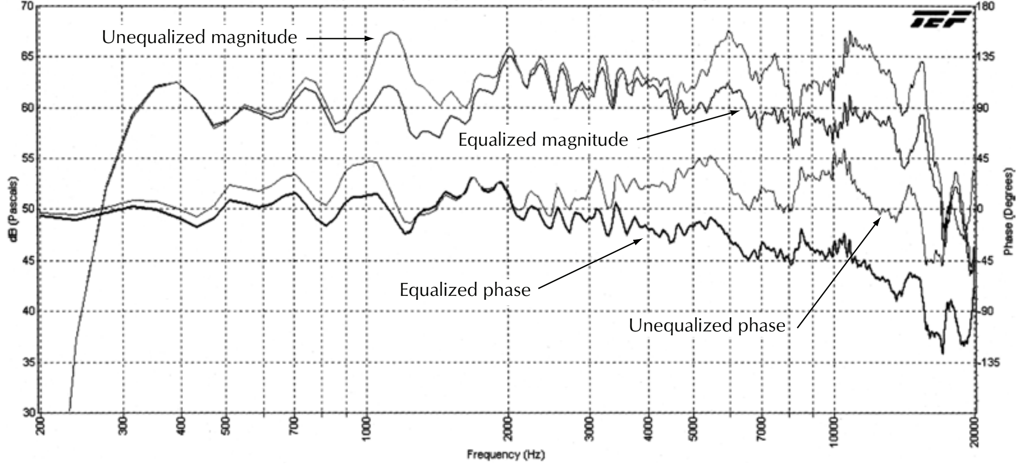


Figure 11-9. Equalized and unequalized transfer functions.

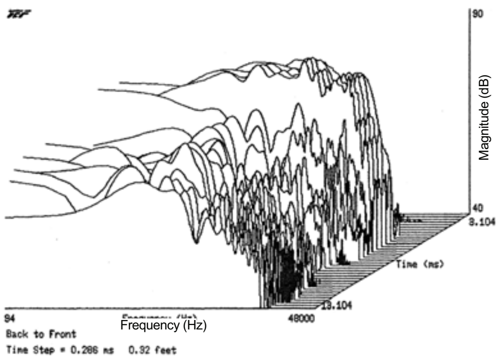


Figure 11-10. Three parameter measurement.

11.5.5 The Nyquist Plot

A Nyquist plot provides simultaneously the real part, the imaginary part, the magnitude, the phase, the frequency and the identification of minimum and non-minimum phase. It is easily one of the most useful frequency domain measurements either electrically or acoustically. Modern analyzers provide cursor read out of the entire plot in addition to identifying non-minimum phase frequencies when they encircle the origin.

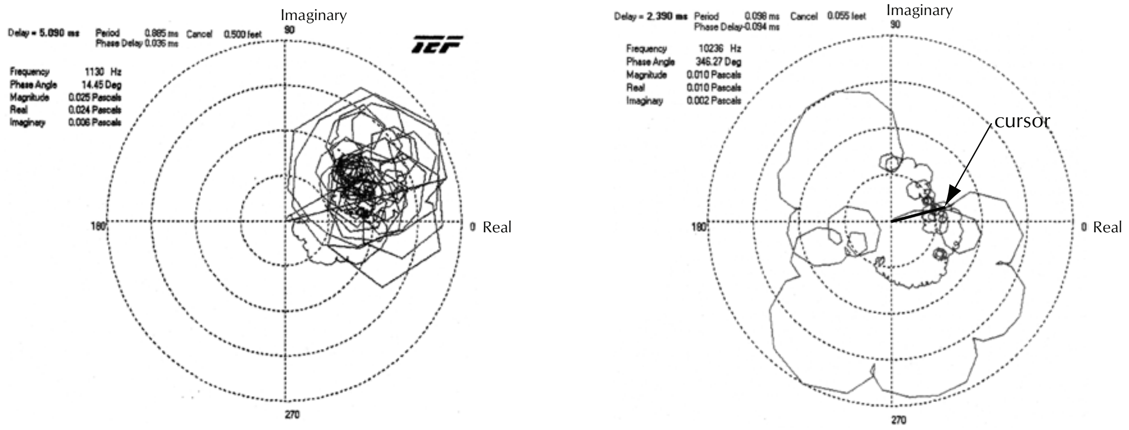
The zero axis is the real component, the 90° axis the imaginary component, and the length from the origin to any chosen frequency on the plot is the magnitude component. Any range of frequency for which the Nyquist plot completely encircles the origin is a range of non-minimum phase behavior. The angle between the zero axis and the cursor set on a given frequency is the phase angle, Figs. 11-11A and 11-11B.

Nyquist Complex Impedance Plots

Many of the points discussed here appeared in an *Audio Magazine* article written by Richard Heyser in June 1984.

The electrical impedance of a loudspeaker is a measure of the amount by which it impedes the passage of current. Impedance is measured by determining the voltage required to pass a fixed amount of current. There is a resistive component and a reactive component to impedance. A loudspeaker can temporarily store some of that energy it gets from the amplifier as well as dissipate the energy in the form of heat and sound. The part that represents dissipation is resistance, the part that represents storage is reactance. The unit of measurement is the ohm. 1.0 A current produces one volt drop across one ohm impedance. If a loudspeaker were a pure resistance load, energy would only pass from the amplifier to the speaker, where it could be converted to heat and sound, but a loudspeaker is not a pure resistor. Loudspeakers store energy and send it back to the amplifier as the amplifier attempts to maintain control of the signal volt.

Heyser's preferred mode of plotting impedance was the Nyquist plot. The Nyquist plot always curls clockwise as it progresses upward in frequency. The plot looks like, and is, circles on circles. The circle form is a fundamental expression of energy



Note: Non minimum phase angle for cursor (dark line) is −346.27°

A. Minimum phase Nyquist plot.

B. Non-minimum phase Nyquist plot.

Figure 11-11. Minimum and non-minimum Nyquist plots.

exchange. As an electromechanical device, the loudspeaker will have a number of impedance resonance modes. A sealed low-frequency driver will have one bass resonance circle, while most vented low-frequency units have two. An epicycle on the low-frequency plot of a sealed system often indicates a poorly sealed enclosure with a leak.

> *Sometimes separate drivers in a system will talk to each other. Acoustic coupling between drivers is always unavoidable, but if improper crossover design allows two or more drivers to carry on simultaneous conversations in the same frequency range, each driver will hear the other talking and show it as a change in impedance. Small extra loops which look like pigtails added to the curve are telltale clues to this interspeaker chitchat.*

Because present-day loudspeakers are designed to produce sound pressure based upon constant voltage applied to the speaker terminals, it would make sense to measure the amount of current that is drawn at this rated voltage. This is just the inverse of impedance. Whereas impedance is a measure of the voltage drop produced by a fixed amount of current, admittance is a measure of the current drawn when a fixed voltage is applied in the simple case of the loudspeaker, admittance is the inverse of impedance.

Heyser preferred admittance measurements which also comes in two parts and are related to dissipation and storage of energy.

> *The part related to dissipation is called conductance, and the part related to storage is called susceptance. The units of measurement for these parts are the inverse of the units of measurement for impedance and are expressed in Siemens.*

Heyser liked to point out that loudspeakers are notoriously nonlinear in their electromechanical properties. Impedance (and admittance) is a function not only of instantaneous drive level, but of the immediate past history of the signal which has been applied to the speaker. They are non-Markovian in their signal handling properties. Clearly one of the most meaningful measurements one can ask in the specification of loudspeakers is a carefully made Nyquist plot of its impedance.

11.5.6 The Polar Envelope Time (PET) *Plots*

The polar envelope time plot allows for any given point of measurement, instant values of:

1. The direction from which the reflection came.
2. The time of travel and distance.
3. The level.

11.5.7 History of Polar Time Measurements

During WWII, Dr. Sidney Bertram developed a Sonar system for submarines named by its users as “Hells Bells.” It consisted of a rotating hydrophone connected to an oscilloscope display through a bank of bandpass filters and associated electronics that displayed direction to target as an angle on the oscilloscope screen and the range to target as variable frequency sound. Close range—low bell-like tones, long range—higher bell-like tones. This system was used to put five U.S. submarines through a dense minefield into the Sea of Japan where they played an effective part in intercepting shipments from the Asian continent.

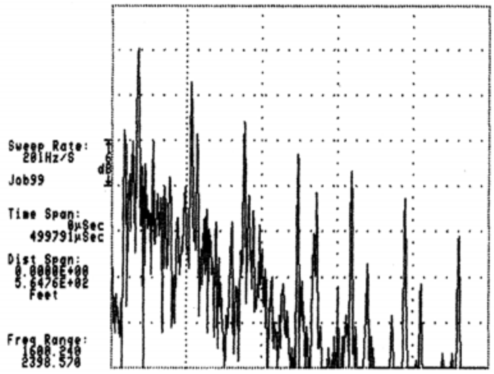
Today’s system uses the input from six directional microphone measurements—forward, right, rear, left, and up and down. Farrel Becker developed PET for use with TEF analysis. It was after Farrel had programmed the software that a report from +30 years after WWII that we read in an IEEE journal gave recognition of Bertram’s work.

The PET measurement is easily one of the most usable measurements ever devised for mapping reflections in architectural spaces. The use of a calibrated cursor gives precise distance, time, and bearing. One sound designer with deep experience in difficult acoustic environments, Deward Timothy of Poll Sound in Salt Lake City, uses the measurement in the orientation of arrays to minimize the detrimental reflections.

Each direction measured produces an individual ETC measurement. These are combined to produce the PET. Samples of each type of display are shown in Figs. 11-12A through 11-12F. To read a PET measurement, identify on the circumference the parameters for that plot, i.e., up, down, etc., or forward, back, left, right, etc.

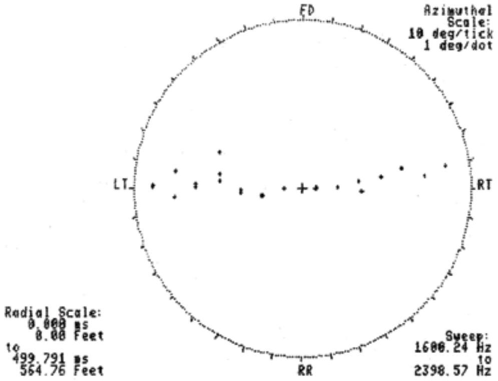
The cursor can be placed on any dot on the screen and its length is the distance, its angle is the bearing, and its magnitude is read on the cursor printout. L_D arrives first and its source is apparent as the shortest distance.

Figs. 11-12A and 11-12B, with the title “Flutter,” are of a severe left-to-right flutter echo. The



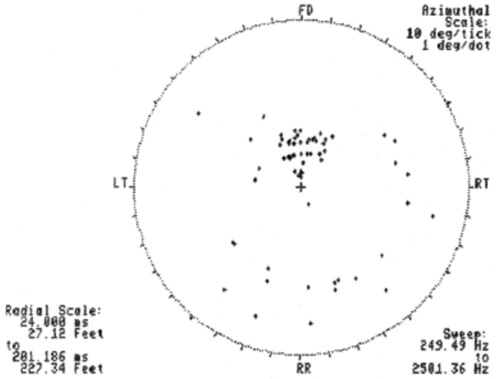
```
Vertical:  6dB/div with base of display at    16.0dB
           0dB is located at  .00002 Pa

Horizontal:        0 microseconds or 0 Feet to
             499791 microseconds or  564.764 Feet
     scale: 1.6177E+02 Feet/inch or  6.3687E+01 Feet/cm.
   1.43155E+5 microseconds/inch or  56360 microseconds/cm.
```

A. Flutter



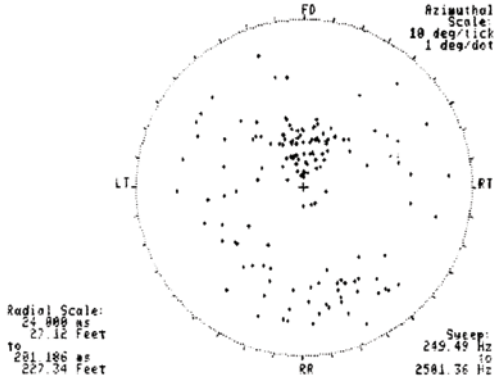
```
Floor Initial Value Below Maximum = 10.0 dB
Floor Final Value Below Maximum   = 39.0 dB
```

B. Flutter



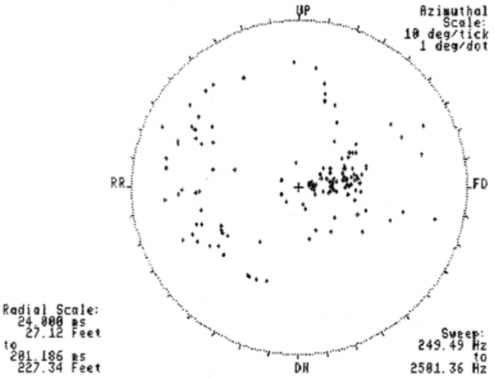
```
Floor Initial Value Below Maximum = 20.0 dB
Floor Final Value Below Maximum   = 20.0 dB
```

C. Room acoustics—front.



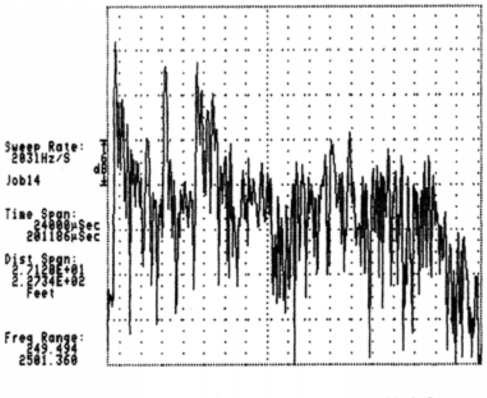
```
Floor Initial Value Below Maximum = 24.0 dB
Floor Final Value Below Maximum   = 24.0 dB
```

D. Room acoustics—front.



```
Floor Initial Value Below Maximum = 24.0 dB
Floor Final Value Below Maximum   = 24.0 dB
```

E. Room acoustics—front.



```
Vertical:  6dB/div with base of display at    11.1dB
           0dB is located at  .00002 Pa

Horizontal:     24000 microseconds or   27.12 Feet to
               201186 microseconds or  227.34 Feet
       scale: 5.7349E+01 Feet/inch or  2.2578E+01 Feet/cm.
        50751 microseconds/inch or  19980 microseconds/cm.
```

F. Conventional ETC.

Figure 11-12. PET measurements.

conventional ETC is one of the four used to make the Polar Time Plot, in this case the left ETC (microphone facing left).

Fig. 11-12C, "Room Acoustics—Front" is from the Intelligibility Workshop. There are three Polar Time Plots. The one with the floor values, first two text lines below the graph, set at 20.0 dB is in the horizontal plane and shows just the strongest reflections. The cluster of reflections in the forward direction, just above the origin, is from the rear wall of the orchestra shell.

Figs. 11-12D and 11-12E have the floor set at 24.0 dB and show more reflections. One is in the horizontal plane and the other is in the median plane. The reflections from the orchestra shell and rear wall of the auditorium are clearly seen in both plots.

Fig. 11-12F is a conventional ETC taken omnidirectionally and from the same location as the polar plots.

11.6 Site Surveys and Noise Criteria Curves

An important test that needs to be made at the site prior to building anything is a noise survey. This can be from a few minutes up to 24 hours. It consists of noise level analysis measurements, NLA, weighted consistent with the existing facts and the expected use of the building. Fig. 11-13A illustrates the variety of data that can be gathered in a one minute example NLA.

Coupled with such measurements should be the established noise criteria desired in order to estimate the required noise isolation the structure must provide, Fig. 11-13B, i.e., the difference between NLA levels and desired criteria.

Once the building is finished, the NC is measured for compliance with the chosen design criteria, Fig. 11-13C. The 2 kHz octave band is the one most often used for the *SNR* figure for $\%AL_{CONS}$ calculation.

A very useful estimator of listener response is shown in Fig. 11-13D where the many variables that help shape human responses are considered and tabulated.

The most commonly encountered violation measured is a too-noisy HVAC system. Often balancing of the HVAC can provide the difference between "fail" and "pass." Because speech intelligibility is directly dependent upon *SNR* failure to specify correct noise criteria, and further failure to measure the violation can doom an otherwise successful sound system installation.

NC curves are plotted in 1⁄1 octave bands. They allow, at a glance, a comparison of the acoustic response at a listener from the sound system to the NC value for the signal-to-noise evaluation. Criteria exist for most common applications.

11.6.1 Constant Percentage Bandwidth Analysis

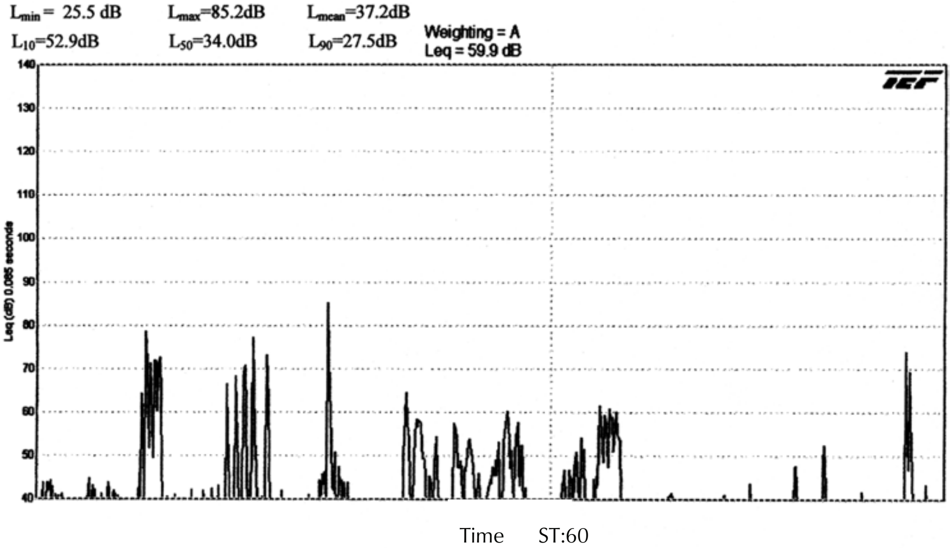
Constant percentage bandwidth analyzers are widely used. One-third of an octave, 23% of center frequency, one-sixth of an octave, 11.5% of center frequency, or one-twelfth of an octave, 5.76% of center frequency filters allow essentially "real time" analysis of non-stationary signals. (See Section 11.9 *Fractional Bandwidth Filter Analyzers.*) One of the authors was instrumental in seeing the first ⅓ octave filter analyzers come to the market in 1968. The term "third octave" is often employed for ⅓ of an octave but is incorrect as it describes a filter for every third octave. See Fig. 11-14A for a ⅙ octave display and Fig. 11-14B for a 1⁄12 octave RTA display.

11.7 An Improper Use of Real Time Analysis

Constant percentage bandwidth filters have absolute widths that increase in direct proportion to the center frequency of the filter. When performing spectrum analysis with instruments based on such filters it is necessary to employ a random noise source whose spectrum has constant energy per octave, i.e., pink noise as opposed to a noise source that has constant energy per unit bandwidth, i.e., white noise.

A system possessing a uniform or flat response on a per unit bandwidth basis that is excited with pink noise will produce a flat display on a constant percentage bandwidth analyzer. Such a system excited with white noise would produce a response that rises at 3 dB/octave on a constant percentage bandwidth analyzer. Therefore, when constant percentage bandwidth analyzers are employed to study the spectra of program material where it is desired to determine the response displayed on a per unit bandwidth basis, it is necessary to precede such an analyzer by a filter that has a response that falls at the rate of 3 dB/octave. Any evaluation of program material without such a device is invalid.

It is the authors' belief that this uncorrected error is why so many professional mixing engineers still use meters and indicators in place of the much more useful real time analyzer. Trained ears didn't agree with the uncorrected visual display. The noise control people made their *criteria* constant percentage bandwidth based, thereby judging relative results. The recording engineers, home hi-fi enthusiasts, and other researchers did not realize the need and therefore failed to compensate for it.



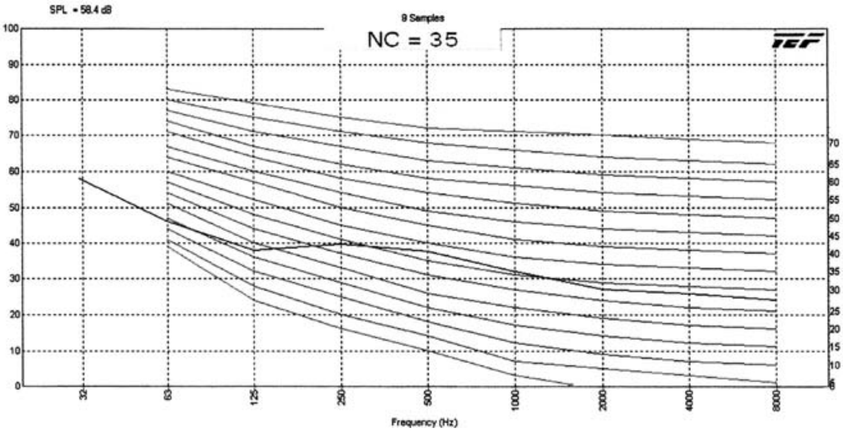
A. NLA sample.

	A Weighted	NC
Residences		
Private homes (rural and suburban)	25-35	20-30
Private homes (urban)	30-40	25-35
Apartment houses, 2- and 3-family units	35-45	30-40
Hotels		
Individual rooms or suites	35-45	30-40
Ball rooms, banquet rooms	35-45	30-40
Halls and corridors, lobbies	40-50	35-45
Garages	45-55	40-50
Kitchens and laundries	45-55	40-50
Hospitals and Clinics		
Private rooms	30-40	25-35
Operating rooms, wards	35-45	30-40
Laboratories, halls and corridors, lobbies and waiting rooms	40-50	35-45
Washrooms and toilets	45-55	40-50
Offices		
Board room	25-35	20-30
Conference rooms	30-40	25-35
Executive office	35-45	30-40
Supervisor office, reception	35-40	30-45
General open offices, drafting rooms	40-55	35-50
Halls and corridors	40-55	35-55
Tabulation and computation	45-65	40-60
Auditoriums and Music Halls		
Concert and opera halls, studios for sound reproduction	25-35	20-25
Legitimate theaters, multipurpose halls	30-40	25-30
Movie theaters, TV audience studios, semi-outdoor amphitheaters, lecture halls, planetarium	35-45	30-35
Lobbies	40-50	35-45

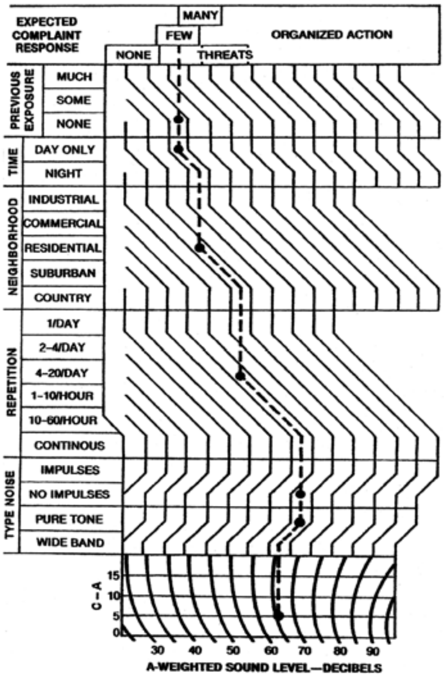
	A Weighted	NC
Churches and Schools		
Sanctuaries	25-35	20-30
Schools and classrooms	35-45	30-40
Laboratories	40-50	35-45
Recreation halls	40-55	35-50
Corridors and halls	40-55	35-50
Kitchens	45-55	40-50
Public Buildings		
Public libraries, museums, court rooms	35-45	30-40
Post offices, general banking areas, lobbies	40-50	35-45
Washrooms and toilets	45-55	40-50
Restaurants, cafeterias, lounges		
Restaurants	40-50	35-45
Cocktail lounges	40-55	35-40
Night clubs	40-50	35-45
Cafeterias	45-55	40-50
Stores retail		
Clothing stores, department stores (upper floors)	40-50	35-45
Department stores (main floor), small retail stores	45-55	40-50
Supermarkets	45-55	40-50
Sports activities—Indoor		
Coliseums	35-45	30-40
Bowling alleys, gymnasiums	40-50	35-45
Swimming pools	45-60	40-55
Transportation (rail, bus, plane)		
Ticket sales offices	35-45	30-40
Lounges and waiting rooms	40-55	35-50

B. Ranges of indoor design goals for air-conditioning system sound control

Figure 11-13. Noise criteria.



C. NC in residence.



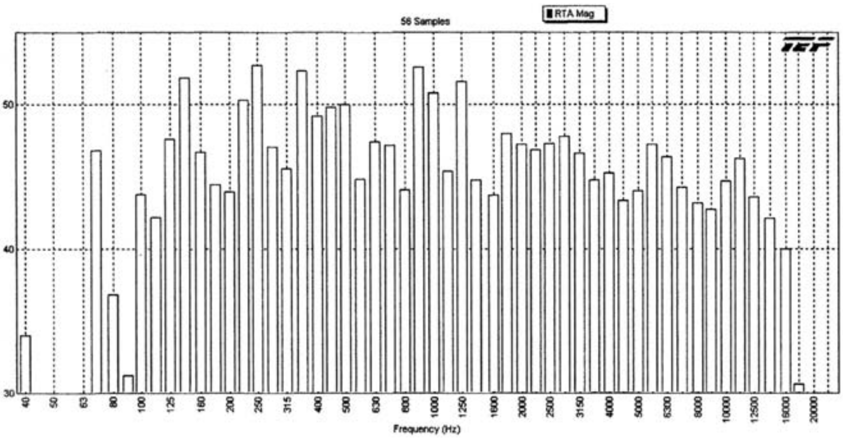
D. Annoyance of neighborhood sound levels.

Figure 11-13. (cont) Noise criteria.

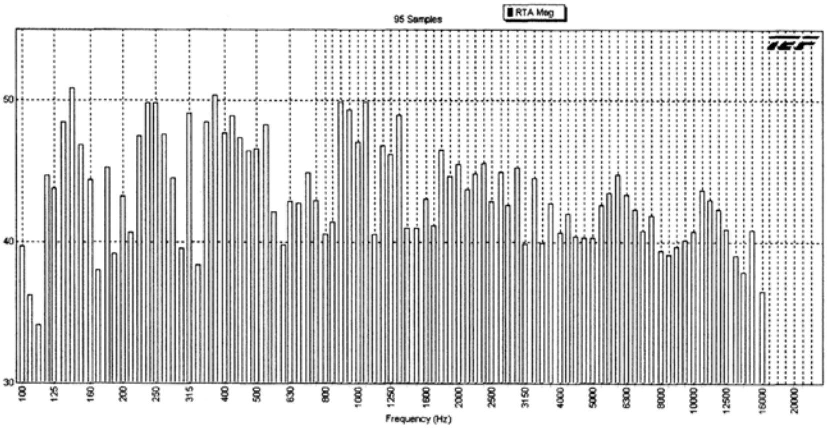
11.8 Evaluation of Listener Response

Measurements that correlate well with listener response are those with a frequency resolution approximating "critical bandwidths." This falls between ⅓ and ⅙ of an octave, Fig. 11-15. One of the authors once walked into a church auditorium to be told by the organist that it was the "deadest" church acoustically he had ever played in, followed a minute later when he met the minister who stated it was "too live" to preach in. Measurements of the reverberation time vs. frequency revealed that something was passing low frequencies from the room but the mid-band frequencies were excessively reverberant. The low frequencies were being diaphragmatically absorbed by the paneling in the ceiling which, when braced, allowed the bass to remain in the room to the satisfaction of the organist. Absorptive treatment on the rear walls and seat cushions controlled the mid-frequency excess to the satisfaction of the minister.

Listener response can vary drastically due to the pinnae response of the individual. For a uniform (flat) frequency response to the ear, these pinnae



A. $^1/_6$ octave RTA display.



B. $^1/_{12}$ octave RTA display.

Figure 11-14. RTA displays.

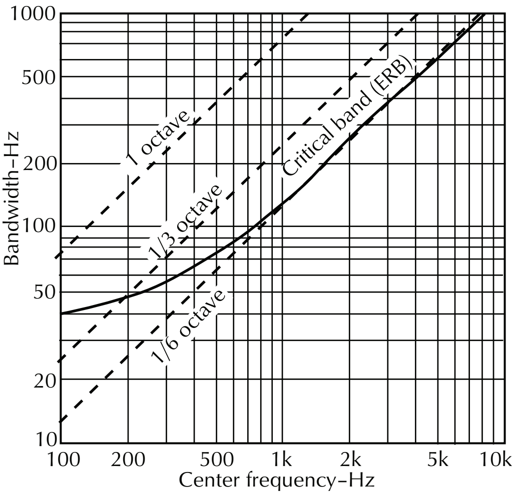


Figure 11-15. A plot of critical bandwidths of the human auditory system compared to constant percentage bandwidths of filter sets commonly used in acoustical measurements.

responses are measured at the individual's eardrum. The combined ear canal resonance and the "comb filtering" of the folds in the pinnae give the resulting responses. Listener response is important because in existing structures, the complaints become microphone positions for measurements that need to be made, Fig. 11-16.

The science and art of measurement begin in the brain of the measurer and the apparatus either confirms the hypothesis or provides the opportunity for serendipity to lead thought in a new direction.

11.9 Fractional Bandwidth Filter Analyzers

Fractional bandwidth filter analyzers are the most commonly used; therefore, we have included their mathematical structure for evaluation of frequency spacing, frequency resolution, and frequency labels from $^1/_1$ to $^1/_{24}$ octave equivalents and $10^{0.3}$ to $10^{0.0125}$ decades. For example, series 80 increments yield

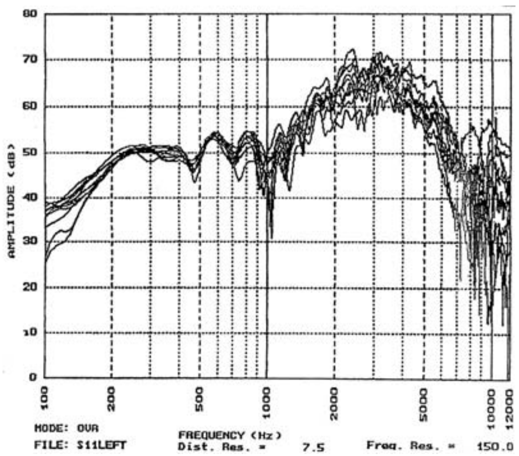


Figure 11-16. Pinnae responses.

series 40 upper and lower crossover frequencies between center frequencies in the 40 series.

The coincidence of ⅓ octaves and ⅒ decades in value has led to wide use of decade filters in both equalizers and analyzers.

1. $2^{1/3}$(1.259921050).
2. $10^{0.1}$(1.258925412).

difference = 0.000995638

$10^{0.1}$ was named the "10" Series because

$$\frac{1}{0.1} = 10 \tag{11-4}$$

There are 10 bands per decade.

Band numbers were $N = 0$ and up ($10^0 = 1.0$). This allowed a very simple equation to be written that would serve for any series and any N in decade intervals. From this equation, any center frequency f_C could then be easily calculated.

$$f_C = 10^{0.1^N} \tag{11-5}$$

simplified to:

$$10^{\frac{N}{series}} \tag{11-6}$$

11.9.1 Series

If the reciprocal of 0.1 is 10 Series, then the ⅟₁ octave equivalent in decades would be 10/3 = $3^{1/3}$ Series (because there are three ⅒ decade f_Cs in the ⅟₁ octave decade equivalent). That makes the decade $10^{0.3}$ for the ⅟₁ octave equivalent. A ⅔ octave decade equivalent would become ⅔(0.3) = $100^{0.2}$. We can now write out for the most used equivalents.

Octave Equiv.	Decade	Series	$N = 1.0$
⅟₁	$10^{0.3}$	3⅓	1.995262
⅔	⅔ × 0.3 = $10^{0.2}$	$1/0.2 = 5$	1.584893
½	½ × 0.3 = $10^{0.15}$	$1/0.15 = 6⅔$	1.412538
⅓	⅓ × 0.3 = $10^{0.1}$	$1/0.1 = 10$	1.258925
⅙	⅙ × 0.3 = $10^{0.05}$	$1/0.05 = 20$	1.220180
⅟₁₂	⅟₁₂ × 0.3 = $10^{0.025}$	$1/0.025 = 40$	1.059254

The series tells how many Ns are in one decade (i.e., $N = 1$ to N for the first decade). The Ns then repeat with only the decimal point moved to the right for each higher decade.

11.9.2 Other Uses

The pass band for filters can be found (for the –3 dB power points) by:

$$f_{xU} = 10^{\frac{N + 0.5}{series}} \tag{11-7}$$

and

$$f_{xL} = 10^{\frac{N - 0.5}{series}} \tag{11-8}$$

$$f_{xU} - f_{xL} = \text{Bandwidth in Hz}$$

where,

f_{xU} is the upper –3 dB point,

f_{xL} is the lower –3 dB point.

The Q of the filter is found by:

$$Q = \frac{f_C}{BW} \tag{11-9}$$

where,

f_C is the center frequency of the filter.

The percent bandwidth is found by calculating the bandwidth for 100 Hz. For example:

$$N = series \times \log f_C \tag{11-10}$$

$$N = 10\log 100 = 20$$

$$\left(10^{\frac{20 + 0.5}{10}}\right) - \left(10^{\frac{20 - 0.5}{10}}\right) = 23.076752\%$$

$$Q = \frac{100}{23.076752}$$

$$= 4.33$$

11.9.3 Useful Tools

To rapidly calculate for any series all f_{xU}, f_C, and f_{xL} you can repeatedly multiply:

$$f_{xU}, f_C, f_{xL} = 10^{\frac{N-0.5}{series}} \qquad (11\text{-}11)$$

The value for $N = ½$ is the first f_{xL}. The next number is $N = 1 f_C$. The next number is f_{xU} for N = 1 and f_{xL} for $N = 2$. Again, the next number is f_C for $N = 2$, etc. Having thus calculated all the f_xs you can extract all the bandwidths.

11.9.4 Using the Decade Exponents

Writing out again the decade exponents we find that by using the ⅒ decade bandwidths we can multiply that bandwidth by 10 times the exponent to obtain the other bandwidths and percentages.

$10^{0.3}$	3 × 23.029 = 69.087
$10^{0.2}$	2 × 23.029 = 46.085
$10^{0.15}$	1.5 × 23.029 = 35.544
$10^{0.1}$	1.0 × 23.029 = 23.029
$10^{0.05}$	0.5 × 23.029 = 11.515
$10^{0.025}$	0.25 × 23.029 = 5.757

For values of Q we can take the 100 Hz/BW for the series 10 and again multiplying the exponents by 10 we can divide them into the series 10 value to obtain all the other Q values.

The difference in power bandwidths expressed in decibels is

$10^{0.3}$	4.342/3.0 = 1.447
$10^{0.2}$	4.342/2.0 = 2.171
$10^{0.15}$	4.342/1.5 = 2.895
$10^{0.1}$	4.342/1.0 = 4.342
$10^{0.05}$	4.342/0.5 = 8.685
$10^{0.025}$	4.342/0.25 = 17.369

$$dB = 10\log\frac{Highest\ frequency\ bandwidth}{Lowest\ frequency\ bandwidth} \qquad (11\text{-}12)$$

Also band 43–13 is 30 bands or 30 dB in the case of series 10.

11.9.5 Decibels

Once again, using series 10 every step in the repeated multiplication of

$$10^{\frac{N-0.5}{10}} \times 10^{\frac{N-0.5}{10}} \times \ldots$$

is a 1.0 dB step on the 20log scale. Every other step (i.e., the f_Cs) is 1.0 dB on the 10log scale.

11.9.6 Label Frequencies

The equations discussed in this article are for the exact frequencies you would use as center frequencies and crossover frequencies for equalizers and analyzers. In actual practice the center frequencies are labeled in a simplified manner. Table 11-1 outlines these labels for the ⅟1, ½, ⅔, ⅓, ⅙, and ⅟12 devices in use today.

Table 11-1. Frequency Labels for Audio Components

⅟12 Octave 40 Series	⅙ Octave 20 Series	⅓ Octave 10 Series	½ Octave 6⅔ Series	⅔ Octave 5 Series	⅟1 Octave 3⅓ Series	Exact Value
1.0	1.0	1.0	1.0	1.0	1.0	1.000000000
1.06						1.059253725
1.12	1.12					1.122018454
1.18						1.188502227
1.25	1.25	1.25				1.258925411
1.32						1.333521431
1.4	1.4		1.4			1.412537543
1.5						1.496235654
1.6	1.6	1.6		1.6		1.584893190
1.7						1.678804015
1.8	1.8					1.778279406
1.9						1.883649085

Table 11-1. (cont) Frequency Labels for Audio Components

1⁄12 Octave 40 Series	⅙ Octave 20 Series	⅓ Octave 10 Series	½ Octave 6⅔ Series	⅔ Octave 5 Series	1⁄1 Octave 3⅓ Series	Exact Value
2.0	2.0	2.0	2.0		2.0	1.995262310
2.12						2.113489034
2.24	2.24					2.238721132
2.36						2.371373698
2.5	2.5	2.5		2.5		2.511886423
2.65						2.660725050
2.8	2.8		2.8			2.818382920
3.0						2.985382606
3.15	3.15	3.15				3.162277646
3.35						3.349654376
3.55	3.55					3.548133875
3.75						3.758374024
4.0	4.0	4.0	4.0	4.0	4.0	3.981071685
4.25						4.216965012
4.5	4.5					4.466835897
4.75						4.731512563
5.0	5.0	5.0				5.011872307
5.3						5.308844410
5.6	5.6		5.6			5.623413217
6.0						5.956621397
6.3	6.3	6.3		6.3		6.309573403
6.7						6.683439130
7.1	7.1					7.079457794
7.5						7.498942039
8.0	8.0	8.0	8.0		8.0	7.943282288
8.5						8.413951352
9.0	9.0					8.912509312
9.5						9.440608688

Table 11-2 is the first decade for all the filters discussed and includes all center frequencies and cross-over frequencies. As a final bonus, from the one exponent, each step is ⅛ dB on the 10log scale and ¼ dB the 20log scale.

1. All even numbers are the 1⁄12 octave center frequencies.
2. Every 4th number is a ⅙ octave center frequency.
3. Every 8th number is ⅓ octave center frequency.
4. Every 12th number is a ½ octave center frequency.
5. Every 16th number is a ⅔ octave center frequency.
6. Every 24th number is a 1⁄1 octave center frequency.
7. All odd numbers are 1⁄12 octave crossover frequencies.
8. Every other 1⁄12 octave center frequency is a ⅙ octave crossover frequency.
9. Every other ⅙ octave center frequency is a ⅓ octave crossover frequency.
10. Every other ⅓ octave center frequency is a ½ octave crossover frequency.
11. Every other ½ octave center frequency is a 1⁄1 octave crossover frequency.
12. ⅔ octave crossover frequencies start at the sixth line and every 12th line thereafter.

11.10 Measuring Electromagnetic Pollution

The proliferation of electromagnetic fields, even in rural areas, such as cross-country power lines, sometimes using exotic new techniques, cell phone towers, wireless Internet providers, powerful weather radars, etc., should be identified and dealt with when designing and installing complex audio and visual systems.

Table 11-2. 80 Series f_Cs

$10^{\left(\frac{0.5}{40}\right)^{N = 1\text{–}20}}$	$10^{\left(\frac{0.5}{40}\right)^{N = 21\text{–}40}}$	$10^{\left(\frac{0.5}{40}\right)^{N = 41\text{–}60}}$	$10^{\left(\frac{0.5}{40}\right)^{N = 61\text{–}80}}$
1.0292005272	1.8302061063	3.2546178350	5.7876198835
1.0592537252	1.8836490895	3.3496543916	5.9566214353
1.0901844924	1.9386526360	3.4474660657	6.1305579215
1.1220184543	1.9952623150	3.5481338923	6.3095734448
1.1547819847	2.0535250265	3.6517412725	6.4938163158
1.1885022274	2.1134890398	3.7583740429	6.6834391757
1.2232071190	2.1752040340	3.8681205463	6.8785991231
1.2589254118	2.2387211386	3.9810717055	7.0794578438
1.2956866975	2.3040929761	4.0973210981	7.2861817451
1.3335214322	2.3713737057	4.2169650343	7.4989420933
1.3724609610	2.4406190680	4.3401026364	7.7179151559
1.4125375446	2.5118864315	4.4668359215	7.9432823472
1.4537843856	2.5852348396	4.5972698853	8.1752303794
1.4962356561	2.6607250598	4.7315125896	8.4139514165
1.5399265261	2.7384196343	4.8696752517	8.6596432336
1.5848931925	2.8183829313	5.0118723363	8.9125093813
1.6311729092	2.9006811987	5.1582216507	9.1727593539
1.6788040181	2.9853826189	5.3088444423	9.4406087629
1.7278259805	3.0725573653	5.4638654988	9.7162795158
1.7782794100	3.1622776602	5.6234132519	10.0000000000

The ability to detect concealed radio frequency transmitters, weak electric fields from wall switches and concealed wiring, poorly grounded wiring, dimmer switches, TV and computers, microwave devices, diathermy apparatus, and other unexpected EMF sources in or near the environment where your system will be installed is vital.

The tri-field ™ meter allows such measurements to be made relatively inexpensively. Their extended range broadband meter has a minimum sensitivity for magnetic fields of one mill gauss and an electric field sensitivity of 10 volts per meter (V/m). Its RF microwave sensitivity is 0.01 kV/m. This handheld direct reading meter can help clarify many otherwise baffling effects.

When doing site surveys sensitivity to high-voltage power lines, RF antennas, substations, in addition to the usual audio noise sources, is an important part of the survey.

Another highly valuable handheld instrument is the GLIT (ground loop impedance tester) which momentarily shorts a power line and measures the loop impedance without causing the breaker to open. These are more costly than the electromagnetic meter but are used much more frequently.

It is surprising that so many people will plug a valuable electronic instrument into a power socket without first testing it to see if it is ac or dc, the proper voltage, proper frequency and properly grounded. In one case, a $5000 analyzer was plugged into a circuit with a dimmer switch on it, which in this case, fortunately only blew the fuse in the analyzer.

While involved in helping provide sound for the American National exhibition in Moscow in 1959, we had only 40 Hz Russian voltage sources and our American amplifiers power transformers drew too much exciting current, after a short period the amplifiers would send up a small mushroom cloud. Inasmuch as this was going to occur in the presence of high officials from both Russia and the U. S. at the opening of the National Exhibition, we advised the security personnel of the possibility, with the result that they gathered around the amplifier and did witness the cloud arise. A spare was immediately available and helped finish the ceremony. An emergency 50 Hz generator was flown in from Berlin to power our pavilion allowing our equipment to work satisfactorily.

These antidotical experiences led me to approach every wall socket with a liberal amount of paranoia. Hopefully the manufacturers of contemporary equipment are more international in their design of protection circuits for their electronics and more tolerant of unexpected power sources.

11.11 Conclusion

Acoustic tests are a mental exercise assisted by measured hints. It's when measurements don't agree with the trained ear that a chance for serendipity is at hand.

Bibliography

A. P. G. Peterson and E. E. Gross, Jr. *Handbook of Noise Measurement*, 7th ed. General Radio Co., Concord, MA, 1972.

B. M. Oliver and J. M. Cage. *Electronic Measurements and Instrumentation*. New York: McGraw-Hill, 1971.

D. Davis. *Acoustical Tests and Measurements.* Indianapolis, Ind.: Howard W. Sams, 1965.

_______. "70 Volt Line Tester During Installation," *Syn-Aud-Con Newsletter*, Vol. 1, No. 3 (1974).

_______. "Real Time Audio Spectrometry," *Record. Eng./Prod.* (1971).

_______. "The Twenty-TDS Workshop," *Syn-Aud-Con Tech Topics*, Vol. 6, No. 13 (1979).

General Radio Co. *Useful Formulas, Tables, Curves for Random Noise.* Tech. Letter, June 1963.

R. C. Heyser. "Acoustical Measurements by Time Delay Spectrometry," *J. Audio Eng. Soc*., Vol. 15 (1967), p. 370.

________. "Breakthrough in Speaker Testing," *Audio* (Nov. 1973), pp. 20-30.

________. "Determination of Loudspeaker Arrival Times: Part I," *J. Audio Eng. Soc*., Vol. 19 (Oct. 1971), pp. 734-743.

________. "Determination of Loudspeaker Arrival Times: Part II," *J. Audio Eng. Soc*., Vol. 19 (Nov. 1971), pp. 829-834.

________. "Determination of Loudspeaker Arrival Times: Part III," *J. Audio Eng. Soc.*, Vol. 19 (Dec. 1971), pp. 902-905.

________. "Loudspeaker Phase Characteristics and Time Delay Distortion: Part 1," *J. Audio Eng. Soc.*, Vol. 17 (1969), p. 30.

________. "Loudspeaker Phase Characteristics and Time Delay Distortion: Part 2," *J. Audio Eng. Soc.*, Vol. 17 (1969), p. 130.

R. Morrison. *Instrumentation Fundamentals and Applications*. New York: Wiley-Interscience, 1984.

W. R. Kundert and A. P. G. Peterson. "Spectrum Analyses of Stationary Noise Signals," *Sound and Vibration* (June 1969).

Chapter 12

Large Room Acoustics

by Don Davis

12.1 What Is a Large Room? 215
  Use of the Sabinian Equations in Large Reverberant Spaces 215
  The Sabine Equation 217
  Specific Versus Statistical 218
  Calculating the Rate of Decay of Reverberant Sound Energy 218
  Improved Reverberation Time Calculations 219
  Importance of Sabine 219
  Limitations of All Acoustic Equations Based on Geometry and Statistics 220
12.2 Levels Defined: Sound Power Level (L_W), Sound Intensity Level (L_I), and Sound Pressure Level (L_P) 220
  Sound Power Level (L_W) 220
  Sound Intensity Level (L_I) 220
  Sound Pressure Level (L_P) 220
12.3 Levels in Enclosed Spaces 221
  Hopkins-Stryker—US and SI 222
  Other Terms Derived from Hopkins-Stryker 222
12.4 Differentiating Between Reverberant Level and Reverberation Time 224
12.5 Evaluation of Signal-to-Noise Ratio, *SNR* 225
12.6 Analyzing Reflections and Their Paths 226
  Sound System Near Regeneration 226
  How Harmful Is High RT_{60}? 226
  Effect of Reverberation on Intelligibility 226
  Variations in the Measurement of Reverberation 227
12.7 Critical Distance 228
  Q Versus for Controlling D_c 229
  Dc Multipliers and Dividers 229
  Theory of *Ma* 230
  The Effect of the *N* Factor 230
  Factors to Watch for in Rooms 231
12.8 Conclusion 233
Bibliography 233

12.1 What Is a Large Room?

Manfred Schroeder has defined a large room frequency (F_L) as the frequency above which a large number of room modes will be excited to vibrate at the source frequency.

$$F_L = K\sqrt{\frac{RT_{60}}{V}} \qquad (12\text{-}1)$$

where,

F_L is the large room frequency in Hz,

K is 2000 in the SI and 11,885 in the U.S.,

RT_{60} is the apparent reverberation time for 60 dB of decay in seconds,

V is the volume of the room in m^3 or ft^3.

If we assume for sound systems:

1. A low frequency limit of 80 Hz for speech systems.
2. A low frequency limit of 30 Hz for music systems.
3. An RT_{60} of 1.6 s approximates the decay time expected for a minimum density sound field, then a large room volume for speech becomes approximately:

$$V = K^2\frac{RT_{60}}{F_L^2}$$

$$= (2000)^2\frac{1.6}{(80)^2}$$

$$= 1000\ m^3$$

or

$$V = K^2\frac{RT_{60}}{F_L^2}$$

$$= (11{,}885)^2\frac{1.6}{(80)^2}$$

$$= 35{,}313\ ft^3$$

and for very wide range music:

$$V = K^2\frac{RT_{60}}{F_L^2}$$

$$= (2000)^2\left(\frac{1.6}{30^2}\right)$$

$$= 7111\ m^3$$

or

$$V = K^2\frac{RT_{60}}{F_L^2}$$

$$= (11{,}885)^2\frac{1.6}{(30)^2}$$

$$= 251{,}116.84\ ft^3$$

Therefore, in this book a large room will be one in which, for speech, the internal volume is 35,000 ft^3 or greater and, for very wide range music, the internal volume is 250,000 ft^3 or greater.

Since F_L is frequency dependent, we can employ our standard audio analysis technique. First divide the audio spectrum into three decades on a linear frequency scale so that we may treat the first decade as a "small room" acoustic problem while approaching the upper two decades as a "large room" design problem.

Bolt, Beranek and Newman have long utilized an almost identical equation with a K for US system of 11,250 which would, by conversion, be 1,893.14 for the SI, see Fig. 12-1. Also, the equation originally used in LEDE control room work yields almost identical results in practical cases.

$$F_L = \frac{3 \times (velocity\ of\ sound)}{Room's\ smallest\ dimension} \qquad (12\text{-}2)$$

F_L is a transition area and should be viewed as such and not as a rigid fixed frequency. The critical frequency (f_c) is synonymous with F_L and both notations are used in small-room literature.

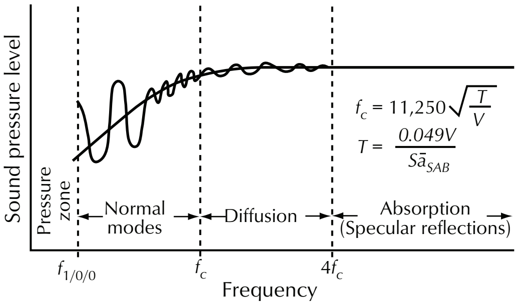


Figure 12-1. Controllers of steady-state room acoustic response. In physically small rooms f_c moves to a frequency of 250 Hz–500 Hz. (*Courtesy Bolt, Beranek, and Newman.*)

12.1.1 Use of the Sabinian Equations in Large Reverberant Spaces

Spaces that qualify as "large rooms" can effectively utilize the myriad of equations based on the original assumptions of Sabine for his reverberation equations. In spaces exceeding these volumes and with an RT_{60} of 1.6 s or greater, we will find mixing homogeneous sound fields of sufficient density to allow accurate engineering estimates of the level of each.

Harvard University found in 1885 that its newly completed Fogg Art Museum had severe acoustical difficulties. President Eliot, head of the university, turned to a young physics professor named Wallace Clement Sabine with the request to "do something" about the problem.

Sabine didn't follow the practices of past generations and hang draperies, place carpets, etc., to "deaden" such a "live" room. Instead he turned from qualitative approaches in finding the solution to a study of the problem on a quantitative basis.

Sabine had at his disposal a number of useful tools to aid in the investigation of the problem. First, there was the troubled lecture room in the Fogg Art Museum. Second, there was nearby Saunders Theater which was considered to have excellent acoustics. Third, the constant-temperature room in the subbasement of the Jefferson Physical Laboratory turned out to be a reverberation chamber. Finally, he had a middle-of-the-road room considered acoustically tolerable, but not much more, in the large lecture room, also in the Jefferson Physical Laboratory building.

With these environments as laboratories, Sabine used the seat cushions from Saunders Theater as his portable absorption, organ pipes and a portable windchest as his sound source, and a stopwatch and his own remarkable hearing as his acoustic test instruments.

After more than two years of intensive research (he often taught classes during the day and did research at night, existing on just a few hours of sleep), Sabine not only had corrected the troubled room by adding the correct amount of acoustical absorption, but as it turned out, he had gathered the raw data for the first important breakthrough in the science of architectural acoustics.

One Saturday evening on the 29th of October, 1898, staring at some of his curves, Sabine called out to his mother (who was living with him at the time), "Mother, it's a hyperbola!" This simple, but inspired, observation took architectural acoustics out of the dark ages of cut-and-try into the sunlight of calculation and measurement.

The insight that came to Sabine, revealing the fundamental relationship between the size of a room and the absorption needed, resulted from his unbelievably precise measurements coupled with his intuitive genius. Thereafter, the reverberation time of a room was calculable prior to construction.

In September, 1975, some 77 years later, W. B. Joyce, in an article entitled, "Sabine's Reverberation Time and Ergodic Auditoriums" in the *Journal of the Acoustical Society of America*, showed the relationship between the second law of thermodynamics and Sabine's equation. This talented Bell Laboratories scientist derived Sabine's equation from a literature search that could have been done at Sabine's time since the necessary thermodynamic concepts were extant by 1895. See box.

Sabine's Reverberation Time and Ergodic Auditoriums

by Wm. B. Joyce

Published: J. Acoust. Soc. Am. *Vol. 58, No. 3, pp. 643-655, September 1975*

Abstract: It is shown in geometrical acoustics that ergodic specular enclosures do exist and that in such auditoriums, but not in general, 4V/S′ is the exact mean directed path length (V is volume and S′ is any part of surface area S). Sabine's expression is then demonstrated to yield the exact reverberation time, provided the enclosure is mixing and provided the inhomogeneous anisotropic surface absorptivity is sufficiently weak. It is further proven that the fundamental form of Sabine's expression cannot be modified so as to become correct for large absorption. In an attempt to reassign credit and reconcile these results with influential findings to the contrary, a short historical account is added. Conditions imposed upon the surface reflectivity—whether the reflectivity be reversible (specular) or other or irreversible (statistical)—by the second law of thermodynamics and by the principle of detailed balance are evaluated. Extensions (e.g. mean length of curved paths in an ergodic auditorium with a thermal gradient) and other applications (electroluminescent diode design) are noted.

In 1929, M. J. O. Strutt considered reverberation by regarding it as a case of free damped vibration of the volume of the air enclosed in a room (this was before computers). The analysis involves the general wave equations, with suitable boundary conditions imposed. Strutt regarded as unsatisfactory the theories which dealt with the paths of separate sound rays (geometric acoustics). The various Eigen Tones or modes of the resonant vibration of the air columns in the room appear in the analysis. This analysis revealed Sabine's law as an asymptotic property toward which the reverberation tends, as the frequency of the (forcing) sound becomes infinitely great compared with the wavelength of the sounds.

Later work at MIT by Philip Morse and Richard Bolt led to the honest but humorous conclusion that "*The practical role of wave acoustics is that it can indicate how to design an enclosure for which geometrical acoustics and statistical acoustics are valid, and in which there is no need of wave acoustics.*"

Fig. 12-2 illustrates the typical measurement setup. Pink noise is emitted by the loudspeaker until a steady state level is produced in the enclosure (i.e., the rate of acoustic energy emission is equal to the the rate of acoustical energy absorption). Then the amplifier is disconnected from the loudspeaker. The microphone signal is fed through a bandpass filter (typically either an octave or ⅓ octave), and the decay rate is observed on the display unit (which may be a digital meter, a graphic level recorder, or an oscilloscope screen). When a graphic level recorder is used, Fig. 12-3 shows how the trace produced is analyzed.

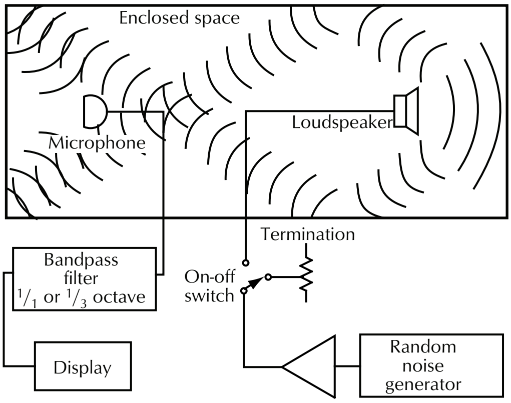


Figure 12-2. Measuring the RT_{60} of an enclosure.

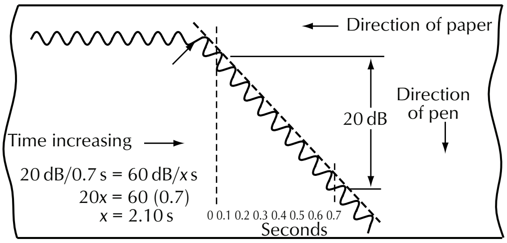


Figure 12-3. Chart recorder method of measuring RT_{60}.

12.1.2 The Sabine Equation

William Joyce at Bell Telephone Labs has demonstrated that the Sabine Equation is fundamental to the decay of energy in a light emitting diode just as it is in an enclosed acoustic space.

The accuracy of this remarkable equation is dependent upon a space being ergodic. There must be a sufficient number of reflections, of a statistical nature, that allow the concept of the mean free path to be meaningful.

In physics, the mean free path is defined as:

$$MFP = \frac{4V}{S} \qquad (12\text{-}3)$$

where,

V is the volume of the space enclosed in some unit of length, l^3,

S is the internal surface area in l^2. This allows any convenient system to be used, i.e., inches, feet, yards, millimeters, centimeters, meters, etc.

The amount of time it takes a sound field to decay to one-one-millionth of its original energy, –60 dB, after the sound source is turned off is called the reverberation time, RT_{60}.

The principle cause of the energy decay is the amount of acoustic absorption available in sabin, $S\bar{a}$. Materials are rated in absorption units (dimensionless $\bar{a}$) that range from 0.0 (totally reflective) to 1.0 (totally absorptive.) Absorption becomes dimensional when multiplied by the area it is on which makes $S\bar{a}$ in l^2 (see Table 10.4 in Chapter 10 *The Acoustic Environment* for absorption coefficients commonly encountered).

We can write an equation that predicts the number of reflections (N) that will occur during a 60 dB decay

$$\begin{aligned} N &= 6\ln(10)\left(\frac{1}{\bar{a}}\right) \\ e^{\ln(10)} &= 10 \\ e^{6\ln(10)} &= 10^6 \end{aligned} \qquad (12\text{-}4)$$

or

$$\begin{aligned} N &= 2.30\log(10^6)\left(\frac{1}{\bar{a}}\right) \\ \ln(10) &= 2.303 \\ e^{2.303\log(10^6)} &= 10^6 \end{aligned} \qquad (12\text{-}5)$$

where,

$\bar{a}$ is the average absorption coefficient of all coefficients present in the enclosure:

$$\bar{a} = \frac{s_1a_1 + s_2a_2 + \ldots + s_na_n}{S_T} \qquad (12\text{-}6)$$

where,

$s_{1,2,3\ldots}$ are the surfaces 1, 2, 3,

$a_{1,2,3\ldots}$ are the coefficients of similarly numbered areas,

S_T is the total boundry surface area.

With these tools at hand it is possible to find the number of reflections per second.

$$RPS = \frac{c}{MFP} \tag{12-7}$$

where,

c is the velocity of sound, light, etc., in l/s.

It then becomes possible to compute the reverberation time with the equation:

$$RT_{60} = \frac{N}{RPS} \tag{12-8}$$

Defined dimensionally:

$$\frac{N\ (\text{Dimensionless})}{\frac{l/s}{\frac{4l^3}{l^2}}} = \text{s} \tag{12-9}$$

For those who work consistently in some preferred dimensional system and prefer to eliminate c from the equation

$$4(2.303)\log(10^6) = 55.26$$

$$RT_{60} = \frac{55.26}{S\bar{a}(c)}$$

$$\frac{55.26}{1130} = 0.049$$

$$\frac{55.26}{344.42} = 0.161$$

results in

$$\begin{aligned} RT_{60} &= \frac{0.049V}{S\bar{a}} \quad \text{U.S. equation} \\ RT_{60} &= \frac{0.161V}{S\bar{a}} \quad \text{SI equation} \end{aligned} \tag{12-10}$$

Example:

Let:

$V = 500{,}000\,\text{ft}^3$.

$S = 42{,}000\,\text{ft}^2$.

$\bar{a} = 0.128$.

then

$$\begin{aligned} MFP &= \frac{4(500{,}000)}{42{,}500} \\ &= 47\text{ ft} \end{aligned}$$

$$\begin{aligned} N &= 2.303\log(10^6)\left(\frac{1}{0.128}\right) \\ &= 108\text{ reflections} \end{aligned}$$

$$\begin{aligned} RPS &= \frac{1130}{47} \\ &= 24\text{ reflections per second} \end{aligned}$$

$$\begin{aligned} RT_{60} &= \frac{108}{24} \\ &= 4.5\text{ s} \end{aligned}$$

12.1.3 Specific Versus Statistical

One hundred eight reflections allow a reasonable statistical sample. When small absorptive spaces such as control rooms in recording studios, small classrooms, etc., are computed the inapplicability of statistical equations becomes apparent because of the low N. Such enclosures do indeed have a finite number of reflections that are best handled by careful Envelope Time Curve (ETC) analysis and specific rather than statistical treatment of the indicated surfaces.

12.1.4 Calculating the Rate of Decay of Reverberant Sound Energy

The classic Sabine equation is

$$RT_{60} = \frac{0.049V}{S\bar{a}} \tag{12-11}$$

Because the reverberation time in seconds is calculated to be the time required for the sound energy to decay to 1/1,000,000th (–60 dB) of its original value prior to switching off the sound source, we can further write decay rate in dB/s = 60 dB/RT_{60} and, because

$$\frac{S\bar{a}}{0.049V} = \frac{1}{RT_{60}} \tag{12-12}$$

we can derive the following direct equation for decay rate in dB/s:

$$\frac{60S\bar{a}}{0.049V} = \frac{60}{RT_{60}} = \frac{1224.5S\bar{a}}{V} \qquad (12\text{-}13)$$

For example, in a church that has an RT_{60} of

$$RT_{60} = \frac{0.049(500{,}000)}{9800} = 2.5\ \text{s}$$

and a decay rate of

$$\text{dB/s} = \frac{1224.5(9800)}{500{,}000} = 24\ \text{dB/s}$$

To check, we can take

$$\frac{60}{2.5} = 24\ \text{dB/s}$$

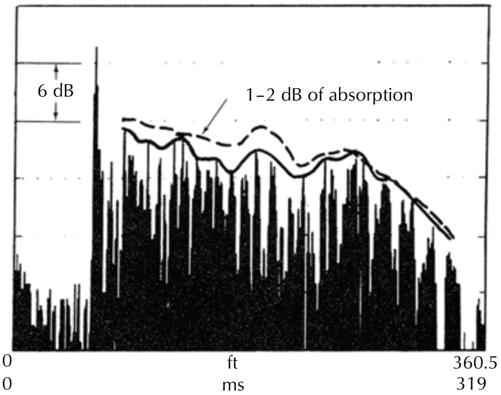
12.1.5 Improved Reverberation Time Calculations

The works of Joyce and Gilbert have increased our appreciation for the fundamental integrity of the original Sabine equation when used in spaces where it can properly be applied. Our work with the TEF analysis process has confirmed that the simple equations that work so well in "live" rooms should simply not be applied in any form in small dead rooms.

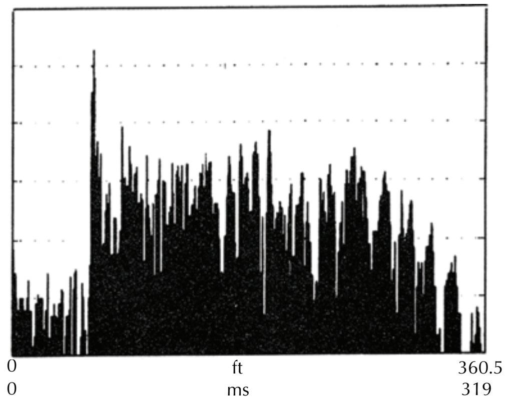
One of the confusions that can arise is that absorption is useful for the control of specular reflections in rooms where no real reverberant sound field exists and application of the statistical formulas is nonsensical. Indeed, application of absorption in these cases is immediately audible, often dramatically so, whereas massive absorption in large "live" rooms that meet the classic criteria do indeed provide a lowering of the statistical reverberant field level but is a much more subtle audible affect. In the "grey areas", we find influences from both approaches and care must be taken to use a sufficiently conservative design approach that allows for "worst case" possibilities.

It has been our experience that the majority of listeners who declare a room as "live" or "dead" do so on the basis of initial signal delay gap and the level of the first reflections and not on the level of the reverberant sound field.

Now that we can view instrumentally the density and spectral uniformity, as well as the changes in frequency with time exhibited by sound fields in real spaces, the true cause and effect relationships exerted by the boundary surfaces of a space will become even more accessible than at present. One aspect that is particularly interesting is the length of time required to see the first effect of the presence of absorption in a large hall, see Figs. 12-4A and 12-4B for two ETCs of before and after absorption.



A. With absorption curtain extended in an auditorium (absorption present), 1 dB–2 dB of absorption is very audible.



B. Auditorium with curtain retracted (without absorption).

Figure 12-4. RT_{60} measurement made with and without absorption.

12.1.6 Importance of Sabine

The work of Wallace Clement Sabine founded the entire field of architectural acoustics and is fundamental to the successful interface of any electroacoustic system to the acoustic environment. A partial list of present day equations directly based on Sabine's work is:

1. Critical distance.
2. Reverberation.
3. Reverberant sound field.
4. Transmission loss.
5. Hopkins-Stryker and its many variations.

6. Articulation loss of consonants.
7. Q relative to the adjustment of direct-to-reverberant ratios.

The genius of the man is apparent, important, and yet relatively unheralded. Encyclopedias rarely mention him. Outside of the field of architectural acoustics, students fail to recognize his name. Wallace Clement Sabine deserves our honored respect and acknowledgment.

12.1.7 Limitations of All Acoustic Equations Based on Geometry and Statistics

It should always be considered that, insofar as the reverberation formulas depend upon statistical averages, they presuppose a complete mixing of sound in the room. In very absorptive rooms, the sound dies away in a few reflections, and the statistical basis of the formulas is weakened. In recent studies done with time energy frequency analysis, typical meeting rooms in hotels have been found in some cases, $RT_{60} = 0.5$ s, to develop no reverberant sound field, whereas in others, $RT_{60} = 0.7$ s, a field barely appears.

Our experience with time-energy-frequency measurements causes us to state unequivocally that recording studio control rooms are not proper subjects for use of classic statistical equations.

In very large rooms, such as the Astrodome and the Superdome, because the sound cannot cross the room many times during a measured reverberation period of a few seconds, the validity of the formula is affected.

12.2 Levels Defined: Sound Power Level (L_W), Sound Intensity Level (L_I), and Sound Pressure Level (L_P)

12.2.1 Sound Power Level (L$_W$)

L_W is the total acoustic power level in dB radiated by a sound source.

$$L_W = 10\log\left(\frac{W_a}{10^{-12}\text{ W}}\right) \tag{12-14}$$

where,
W_a is the acoustic watts,
10^{-12} W is the specified reference.

For an output of 1.0 W we can write

$$\begin{aligned} L_W &= 10\log(W_a) + 120\text{ dB} \\ &= 10\log(1.0\text{ W}) + 120\text{ dB} \\ &= 120\text{ dB} \end{aligned}$$

This means that a device radiating a total acoustic power of 1.0 W will have an $L_W = 120$ dB regardless of radius, r, from the source or how confined a directivity factor, Q, happens to be.

12.2.2 Sound Intensity Level (L$_I$)

If we imagine a sphere with a surface area A of 1.0 m^2, the radius r becomes

$$\begin{aligned} r &= \sqrt{\frac{A}{4\pi}} \\ &= 0.282\text{ m} \end{aligned}$$

The 1.0 W radiating from an omnidirectional point source through a spherical surface area of 1.0 m^2 would have a sound intensity level of

$$L_I = 10\log\left(\frac{\dfrac{W_a}{\text{m}^2}}{\dfrac{10^{-12}\text{ W}}{\text{m}^2}}\right) \tag{12-15}$$

Which can again be written as

$$\begin{aligned} L_I &= 10\log\left(\frac{W_a}{\text{m}^2}\right) + 120 \\ &= 10\log\left(\frac{1.0\text{W}}{\text{m}^2}\right) + 120 \\ &= 120\text{ dB} \end{aligned}$$

We would also find at the surface of our imaginary sphere a sound intensity of 1 W/m^2.

12.2.3 Sound Pressure Level (L$_P$)

The root mean square acoustic pressure is given by

$$\begin{aligned} P_{RMS} &= \sqrt{I_a \rho c} \\ &= \sqrt{1.0\frac{\text{W}}{\text{m}^2}(400)} \\ &= 20\text{ Pa} \end{aligned}$$

where,
I_a is the acoustic intensity in W/m^2,

ρc is the specific acoustic resistance of air. ρ is expressed in kg/m^3 and c is the sound speed in m/s. ρc has a value of 400.

There is a sound pressure reference value of 20 micropascals (µPa) = 0.00002 Pa.

$$L_P \text{ (sound pressure level)} = 20\log\left(\frac{P_{rms}}{0.00002\ \text{Pa}}\right)$$

or

$$20\log\left(\frac{P_{rms}}{1\ \text{Pa}}\right) + 94\ \text{dB}$$

and for 20 Pa

$$L_p = 20\log(20) + 94\ \text{dB}$$
$$= 120\ \text{dB}$$

At a radius of 0.282 m, a sphere has a surface area of 1.0 m^2 and 1.0 acoustic watt radiating through that surface area produces at that surface an

L_W = 120 dB,
L_I = 120 dB,
L_P = 120 dB.

This means only that these three parameters are numerically identical.

Next, imagine a hemisphere with a radius of 0.282 m. The source is radiating 1.0 acoustic watt; therefore, the L_W = 120 dB. The surface area is now 0.5 m^2 and the 1.0 W now produces an intensity of 1.0 W/0.5 m^2 or 2.0 W/m^2. This results in L_I = 123 dB and an L_P = 123 dB. The difference between L_I and L_W is called the directivity index, D_I in dB.

$$10^{\left(\frac{D_I}{10}\right)} = Q$$

$$10^{\left(\frac{3.01}{10}\right)} = 2$$

The directivity factor, Q, describes the increase in power per unit of area that results from confining available power to a smaller area. The comparison is between the Q confined area versus that over the spherical power per unit of area. Today Qs of 50+ are available from devices that cover 1⁄50 of a spherical surface, thus simultaneously achieving controlled coverage of an audience area and supplying that area with an L_P that required 1⁄50 the power an omnidirectional device would require for the same L_P, thus yielding a +17 dB advantage (20log50⁄1 = 16.99 dB).

Finally, increase the radius of the original sphere by a factor of 2 such that r = (2)(0.282) m.

$$A = 4\pi r^2$$
$$= 4\pi[2(0.2821)]^2$$
$$= 4\ \text{m}^2$$

The result in a surface area of 4 m^2 and 1.0 acoustic watt now produces:

L_W = 120 dB.
L_I = 114 dB.
L_P = 114 dB.

Or for each doubling of distances we drop 6 dB in level.

$20\log(D_1/D_2)$ = Change in level of L_P

$20\log(0.282) - 20\log(0.564) = -6.02$ dB

The rate of change in level is a consequence of the inverse square law that governs radiation from point sources.

A given L_W is independent of both distance and area covered. We can state that L_I and L_P each vary with both distance, r, and directivity, Q as:

10log Q or 20log r

$$L_I \text{ or } L_P = L_W + 10\log Q \quad (12\text{-}16)$$

$$L_I \text{ or } L_P = L_W - 20\log\left(\frac{r}{0.282\ \text{m}}\right) \quad (12\text{-}17)$$

where,

r is greater than 0.282 m (0.925 ft).

As will be seen further on we can use these identities to predict efficiencies, power, and pressure relationships.

12.3 Levels in Enclosed Spaces

When a loudspeaker radiates sound into an enclosure, its acoustic performance as determined under open air conditions is modified by the acoustic properties of the space, but the total power L_W radiated is essentially unchanged. When a sound source is turned on in a room, the energy spreads from the source and then strikes the various wall surfaces, S, where it is partially absorbed, a, and partially reflected $1 - a$ to other surfaces which, in turn, absorb and reflect. This process continues until the energy in the room reaches a steady value, i.e., when the rate of energy absorption by the surfaces and in the air becomes equal to the rate of energy emission by the source.

This energy is made up of the total reverberant energy, L_R, assumed uniform in distribution, and the total direct energy L_D. This division is expressed as

$$L_D = L_W + 10\log\left(\frac{Q}{4\pi r^2}\right) + 10.5 \text{ U.S.} \qquad (12\text{-}18)$$

and

$$L_R = L_W + 10\log\left(\frac{4}{S\bar{a}}\right) + 10.5 \text{ U.S.} \qquad (12\text{-}19)$$

Interestingly $4V/S$ considered as a radius to a sphere suggests that the volume of such a sphere and the room volume, for all rooms of reasonable proportions, will be nearly equal, thus allowing a simplification of terms. From these relatively simple relationships it becomes evident that L_W and $S\bar{a}$ determine the reverberant levels and that L_W, Q, and r^2 determine the direct sound level at any given position. In fact the design goal in most cases is to insure that any listener receives at the least an $L_D = L_R$ and at the best that $L_D \geq L_R$.

12.3.1 Hopkins-Stryker—US and SI

The interplay of directivity factor, Q, distance from the sound source to the listener, D_2, the total acoustic absorption in the room, $S\bar{a}$, and the expected sound pressure level, L_p, at the listener, can all be combined in a single equation called the Hopkins-Stryker.

$$L_P = L_W + 10\log\left(\frac{Q}{4\pi D_2^{\,2}} + \frac{4}{S\bar{a}}\right) + 10.5 \text{ US} \qquad (12\text{-}20)$$

$$L_P = L_W + 10\log\left(\frac{Q}{4\pi D_2^{\,2}} + \frac{4}{S\bar{a}}\right) + 0.2 \text{ SI} \qquad (12\text{-}21)$$

Fig. 12-5A illustrates the levels for the direct sound pressure level, L_D, with the distance, the reverberant sound level, L_R, and the total sound level, L_T, for a $Q = 45$, D_2 from 10 ft to 1000 ft, a total room absorption of 5000 ft², and a sound power level, L_P, of 105.6 dB.

Fig. 12-5B is a table of log multiplier and ratio exponents for curves that are neither inverse square law, or totally reverberant, but rather fall between −2 dB per doubling of distance to −5 dB per doubling of distance i.e., for 3 dB per doubling of distance beyond D_c use:

$$9.966\log D_2$$

or

$$\log D_2^{\,0.996}$$

Finally, Fig. 12-5C shows a fully implemented set of modifiers and multipliers for L_D, L_T, −2 dB/doubling of distance, −3 dB/doubling of distance, −4 dB/doubling of distance, and −5 dB/doubling of distance.

Fig. 12-5C is the same case as Fig. 12-5A with the modifiers added of $N = 1.0$, $M_e = 1.5$, and $M_a = 5.0$ where the Hopkins-Stryker equation becomes

$$L_P = L_W + 10\log\left[\frac{Q(Me)}{4\pi D_2^{\,2}} + \frac{4N}{S\bar{a}Ma}\right] + 10.5 \qquad (12\text{-}22)$$

When knowledge of the dB per doubling of distance has been ascertained either by experience or by measurement, the appropriate multiplier or exponents can be inserted into the calculation. Note that these modifications apply only to distances beyond critical distance, see vertical axis notation on Fig. 12-5B.

For those who would like to use solid angle data in place of Q, the Hopkins-Stryker becomes

$$L_T = L_W + 10\log\left[\left(\frac{1}{sr \times D_2^{\,2}}\right) + \left(\frac{4}{S\bar{a}}\right)\right] + 0.2 \qquad (12\text{-}23)$$

where,

L_T is the sound pressure level,

L_W is the sound power level,

sr is the solid angle in steradians,

r is the distance in meters,

$S\bar{a}$ is the total absorption in sabins meters squared.

Remember that the solid angle in steradians = $4\pi/Q$.

12.3.2 Other Terms Derived from Hopkins-Stryker

For those unfamiliar with these terms, the following definitions are useful.

$$D_I = 10\log\left(\frac{4\pi}{sr}\right) \qquad (12\text{-}24)$$

$$sr = \frac{4\pi}{Q} \qquad (12\text{-}25)$$

$L_{EIA} := 71.9$ dB $\quad Q := 45 \quad W := 0.1 \quad D_x := 90 \quad K := 10.5$ (U.S.) (0.2 =S.I.)

$Lw := L_{EIA} - 10\cdot\log(Q) + 10\cdot\log(W) + 60.22 \quad Lw = 105.6$ dB $\quad Wa := 10^{\left(\frac{Lw}{10}\right)}\cdot 10^{-12}$ *

$a_{aud} := 0.8 \quad Ma := \frac{1}{1 - a_{aud}}$ * $\quad Ma = 5 \quad Me := 1.0 \quad Wa = 0.036$

$D := 4..1000$ ft or m $\quad Sa := 5000$ ft² or m²

$L_{dx} := Lw + 10\cdot\log\left[\frac{Q\cdot Me}{4\cdot\pi\cdot\left(D_x\right)^2}\right] + K$ * $\quad f(D) := Lw + 10\cdot\log\left(\frac{Q\cdot Me}{4\cdot\pi\cdot D^2} + \frac{4}{Sa\cdot Ma}\right) + K$

$D_c := 0.141\cdot\sqrt{Q\cdot Sa\cdot Ma\cdot Me} = 149.6$ ft or m $\quad \%eff := \left[10^{\frac{\left[\left(Lw + 10\cdot\log\left(\frac{1}{W}\right)\right) - 120\right]}{10}}\right]\cdot 100$ *

$L_{30} := Lw + 10\cdot\log\left(\frac{Q\cdot Me}{4\cdot\pi\cdot 900}\right) + K$ * $\quad L_r := Lw + 10\cdot\log\left(\frac{4}{Sa\cdot Ma}\right) + K$

$L_{30} = 92.1$ dB $\quad L_{dx} = 82.5$ dB $\quad L_r = 78.1$ dB $\quad \%eff = 36.2$ %

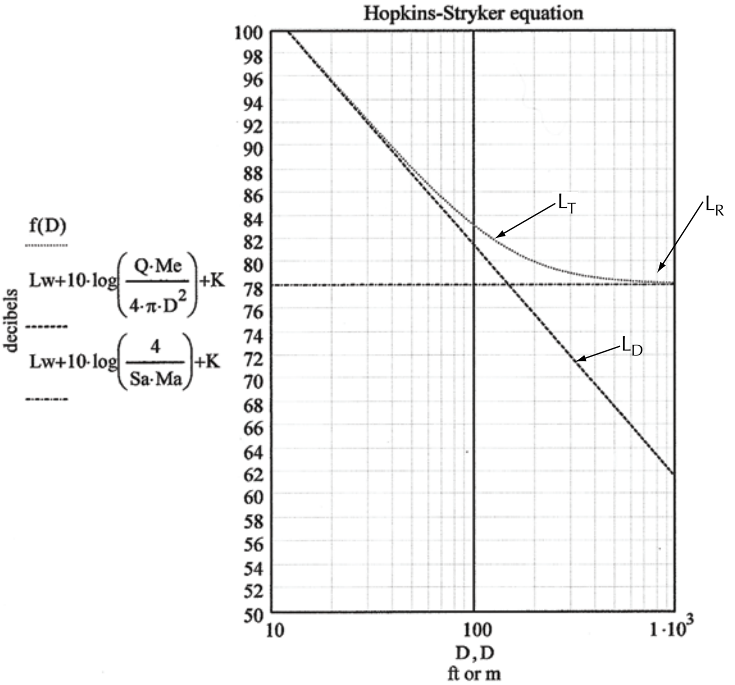


Figure 12-5A. MathCAD program of Hopkins-Stryker plot for system design parameters shown above.

$$Q = \frac{4\pi}{sr} \tag{12-26}$$

$$D_c = 0.5\sqrt{\frac{S\bar{a}}{sr}} \tag{12-27}$$

$$\text{Radians} = Degrees\left(\frac{\pi}{180}\right) \tag{12-28}$$

$$\text{Degrees} = Radians\left(\frac{180}{\pi}\right) \tag{12-29}$$

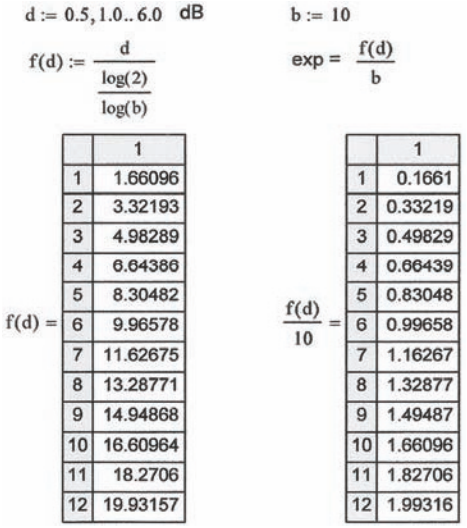
One advantage of using *sr* is that they can be summed to obtain a total *sr* for two adjoining areas. That this has been widely overlooked is apparent in the book, *Units, Dimensional Analysis and Physical Similarity* by D. S. Massey:

> *The magnitude of solid angle may be related to that of plane angle to give the dimensional formula (A^2). However, as the results seem to have no practical use, we shall not discuss further the dimensional formula of solid angle.*

Logarithm Modifiers and Exponents for use in Hopkins-Stryker Equations in 0.5 dB Steps for Changes in Level per Doubling of Distance

Example

For 3.5 dB change for each time the distance is doubled
11.62675log(2) = 3.5dB or 10log($2^{1.16267}$) = 3.5 dB



$d := 0.5, 1.0 .. 6.0$ dB

$f(d) := \frac{d}{\frac{\log(2)}{\log(b)}}$

$b := 10$

$\exp = \frac{f(d)}{b}$

f(d) =

	1
1	1.66096
2	3.32193
3	4.98289
4	6.64386
5	8.30482
6	9.96578
7	11.62675
8	13.28771
9	14.94868
10	16.60964
11	18.2706
12	19.93157

$\frac{f(d)}{10}$ =

	1
1	0.1661
2	0.33219
3	0.49829
4	0.66439
5	0.83048
6	0.99658
7	1.16267
8	1.32877
9	1.49487
10	1.66096
11	1.82706
12	1.99316

Figure 12-5B. MathCAD program of Hopkins-Stryker plot for system design parameters in Figure 5A.

It would seem that radar, sonar, and acoustics are not practical applications, in this author's mind. The book, nevertheless, is excellent and worthy of study.

The direct sound level alone can be expressed as

$$L_D = L_W + 10\log\left(\frac{Q}{4\pi r^2}\right) + 10.5 \text{ US} \qquad (12\text{-}30)$$

$$L_D = L_W + 10\log\left(\frac{1}{(sr)(r^2)}\right) + 0.2 \text{ SI} \qquad (12\text{-}31)$$

and the reverberant sound level, L_R, from

$$L_R = L_W + 10\log\left(\frac{4}{S\bar{a}}\right) + 10.5 \text{ US} \qquad (12\text{-}32)$$

$$L_R = L_W + 10\log\left(\frac{4}{S\bar{a}}\right) + 0.2 \text{ SI} \qquad (12\text{-}33)$$

Finally, critical distance, D_c, is obtained from the fact that when $r = D_c$

$$\frac{Q}{4\pi D_c^2} = \frac{4}{S\bar{a}}$$

$$D_c = \sqrt{\frac{QS\bar{a}}{16\pi}} \quad \text{both in US and SI} \qquad (12\text{-}34)$$

$$= 0.141\sqrt{QS\bar{a}}$$

$$= 0.5\sqrt{\frac{S\bar{a}}{sr}}$$

At D_c the two levels, L_R and L_D are equal; therefore L_T will be +3 dB higher. For a sound system in an enclosed space it is highly desirable to keep as many listeners as possible at or closer to the sound source than D_c.

Bringing the sound source closer to the listener raises the L_D and often, because of being physically closer, inverse square law allows lower power, thus lower L_R. Also raising the directivity factor, Q, of the sound source results in higher L_D for lower power also. Raising the number of sabins in the space lowers L_R.

The accuracy of these equations will get you into the "ball park." They will bring you into the right order of magnitude. Ideally when the space already exists, testing with a sound source of the indicated Q allows much more exact numbers to be computed.

Prediction of sound levels outdoors and in very small non-reverberant spaces will follow inverse square law quite accurately.

Extremely large sporting arenas, domes, etc., have *l*s of such length as to again produce a low number of reflections.

12.4 Differentiating Between Reverberant Level and Reverberation Time

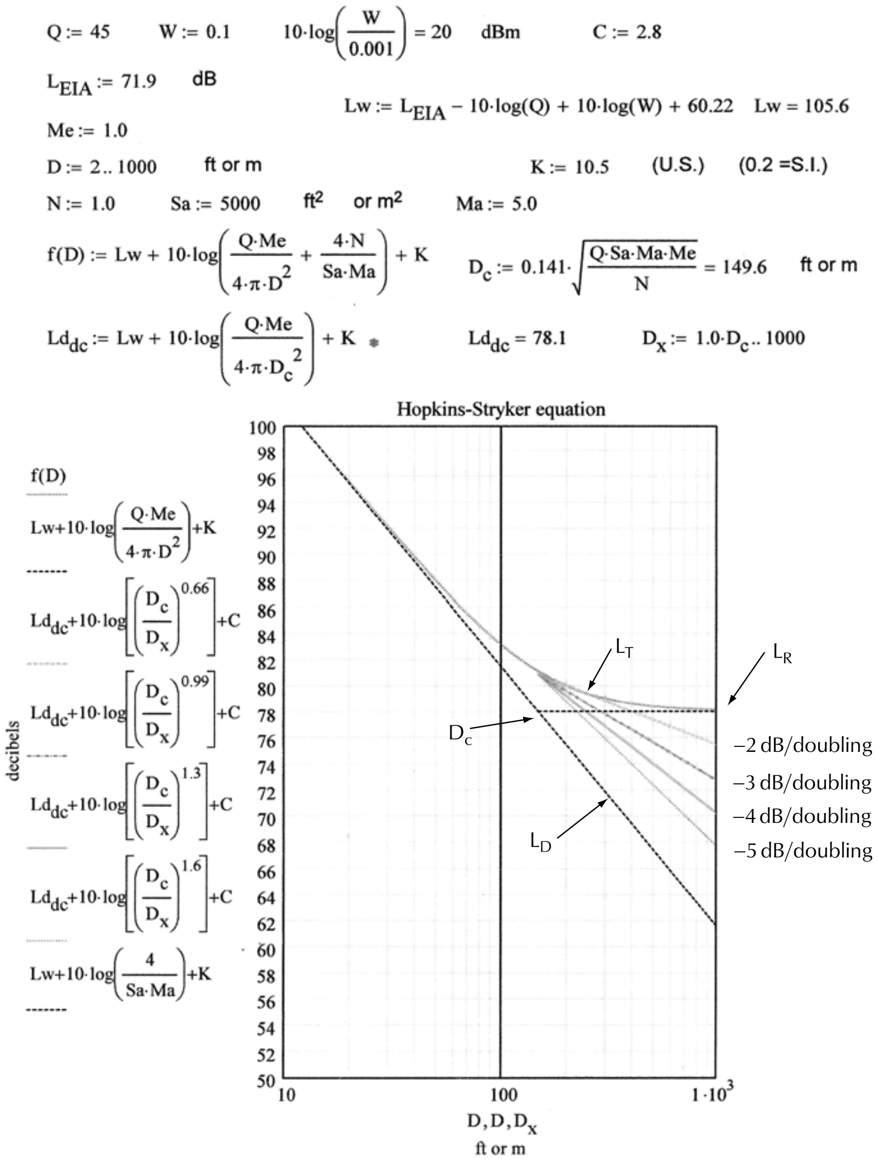
Thanks to present day measurement techniques both L_T and L_D are readily accessed.

When non-uniform reverberant sound fields are encountered, measurements made at repeated doubling of distance beyond critical distance, D_c, allows accessing the Hopkins-Stryker modifier for plotting level versus distance.

$$L_P = L_W + 10\log\left(\frac{Q}{4\pi D_x^2} - \frac{4}{S\bar{a}}\right) + \qquad (12\text{-}35)$$

$$10\log\left(\left(\left[\left(\frac{D_c}{D_x}\right)^{\exp}\right]\right) + K\right)$$

where,



$Q := 45 \quad W := 0.1 \quad 10\cdot\log\left(\frac{W}{0.001}\right) = 20$ dBm $\quad C := 2.8$

$L_{EIA} := 71.9$ dB

$Lw := L_{EIA} - 10\cdot\log(Q) + 10\cdot\log(W) + 60.22 \quad Lw = 105.6$

$Me := 1.0$

$D := 2..1000$ ft or m $\quad K := 10.5$ (U.S.) (0.2 =S.I.)

$N := 1.0 \quad Sa := 5000$ ft² or m² $\quad Ma := 5.0$

$f(D) := Lw + 10\cdot\log\left(\frac{Q\cdot Me}{4\cdot\pi\cdot D^2} + \frac{4\cdot N}{Sa\cdot Ma}\right) + K \qquad D_c := 0.141\cdot\sqrt{\frac{Q\cdot Sa\cdot Ma\cdot Me}{N}} = 149.6$ ft or m

$Ld_{dc} := Lw + 10\cdot\log\left(\frac{Q\cdot Me}{4\cdot\pi\cdot D_c^{\ 2}}\right) + K \qquad Ld_{dc} = 78.1 \qquad D_x := 1.0\cdot D_c..1000$

Hopkins-Stryker equation
f(D)
Lw+10·log(Q·Me/(4·π·D²))+K
Ld_dc+10·log[(Dc/Dx)^0.66]+C
Ld_dc+10·log[(Dc/Dx)^0.99]+C
Ld_dc+10·log[(Dc/Dx)^1.3]+C
Ld_dc+10·log[(Dc/Dx)^1.6]+C
Lw+10·log(4/(Sa·Ma))+K
decibels
100
98
96
94
92
90
88
86
84
82
80
78
76
74
72
70
68
66
64
62
60
58
56
54
52
50
10
100
1·10³
D,D,Dx
ft or m
LT
LR
Dc
LD
−2 dB/doubling
−3 dB/doubling
−4 dB/doubling
−5 dB/doubling

Figure 12-5C. MathCAD program depicting Hopkins-Stryker plots for varying values of attenuation with distance doubling for system design parameters in Figure 12-5B.

$$\exp = \left(\frac{\text{dB/Doubling of distance}}{\frac{\log 2}{10}}\right),$$

L_W is the sound power level in dB,
D_x is any distance l beyond D_c,
$K = 10.5$ US, 0.2 SI.

12.5 Evaluation of Signal-to-Noise Ratio, *SNR*

In my experience you can very seldom turn room noise off, therefore the first measurements are those of room noise levels at various positions with the sound system on but inactive. Subsequent measurements of total level, direct level, and reverberant level can then be corrected for the noise level contribution in those instances where the noise level makes a significant contribution.

Having obtained L_T, L_D, and L_R at any given point, it becomes possible to separate the contribution of the noise level, L_N, by using the measurement of L_T with the noise "on" and then with the noise "off." $L_T - L_N$ is the signal-to-noise ratio expressed in dB. Noise "on" refers to lighting, HVAC, and other man-made noise sources.

In listening to an auditorium where measurements are to be made, always attempt an aural estimate of each of these levels at differing locations at both less than the expected D_c and well beyond D_c. Good sound system design practice tries to minimize the reverberant level while recognizing that changing the reverberation time is a room treatment problem. When a 25 dB *SNR* at 2 kHz is unattainable, recommend the job to a competitor. (See Chapter 11 *Audio and Acoustic Measurements*, for detailed *SNR* analysis.)

12.6 Analyzing Reflections and Their Paths

In sound system analysis, sets of linearly spaced notches and peaks in the amplitude response are called "comb filters." These are the result of a reflection or reflections converging with the direct sound from the desired sound source. Even a low resolution ⅓ octave real-time analyzer can see the lower frequency notch. The reflective path distance, *rp*, can be ascertained from:

$$rp = \frac{0.5c}{F_n} \tag{12-36}$$

where,

c is the velocity of sound. The velocity unit, l/s, can be in ft, inches, mm, m, etc. The reflection distance is in the units chosen,

F_n is the frequency of the first notch.

Because comb filters are spaced linearly, NR_n or NF_p (first peak frequency), this allows the computation of all higher frequency comb filters. For example, a reflection 0.5 mm such as the spacing of a Pressure Zone Microphone, PZM, capsule above its plate results in a first notch at

$$\begin{aligned} F_n &= \frac{0.5c}{diff} \\ &= \frac{0.5 \times (344 \times 10^3)}{0.5\ \text{mm}} \\ &= 344{,}420\ \text{Hz} \end{aligned}$$

which could be disregarded in normal use. If, on the other hand, a low pass filter at 20 kHz, for anti-aliasing purposes, could be begun acoustically near the bandpass by

$$\begin{aligned} \text{Spacing} &= \frac{0.5c}{diff} \\ &= \frac{0.5 \times (13{,}560)}{20{,}000\ \text{Hz}} \\ &= 0.34\ \text{inches} \end{aligned}$$

Figs. 12-6A and 12-6B illustrate the usefulness of comb filters as well as how to analyze harmful ones.

12.6.1 Sound System Near Regeneration

A sound system operating too near positive acoustic feedback amplifies the room's natural reverberation time by many times (as many as 4 to 5 times). Be sure that your subjective judgment of the space has not been influenced by the presence of a malfunctioning sound system.

See Fig. 12-7 from the work of William Snow of Bell Labs. Note that the decay rate changes from 92 dB/s to 22 dB/s as the sound system is brought near sing point (feedback), a 4.2 magnification factor.

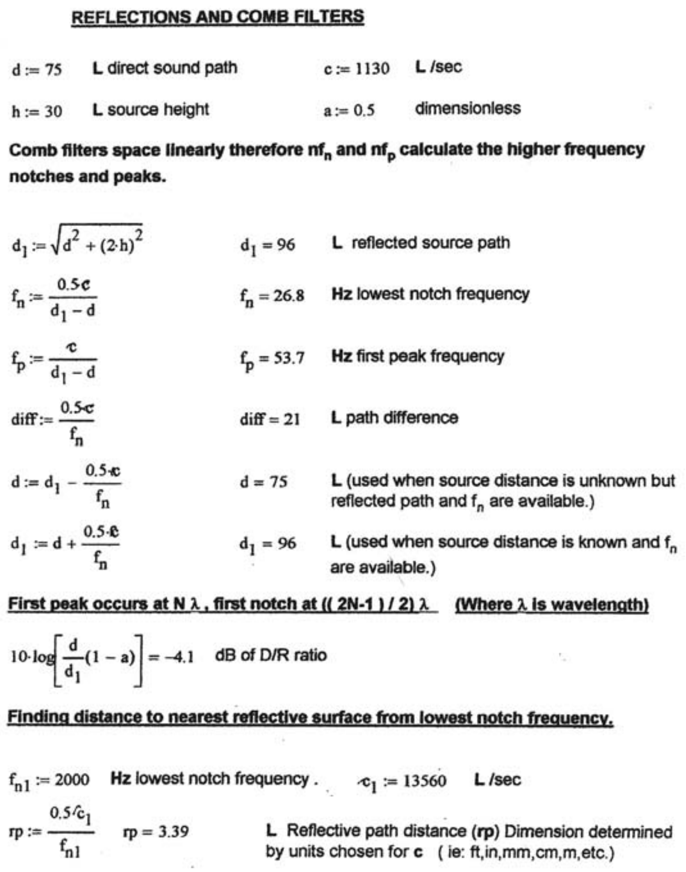
12.6.2 How Harmful Is High RT_{60}*?*

How harmful is a high RT_{60}? In a truly diffuse acoustic field it is surprisingly not harmful to "live" speech until reverberation times around 3 s to 4 s at 2 kHz are reached. The trouble in many "reverberant" spaces is that it is focused energy returns over long path lengths that are the culprit and not the length of time they take to decay. This can usually be demonstrated by using a large sheet of Sonex (4 ft × 5 ft) and circling the listener while a talker speaks from the podium. When the Sonex passes between the focused energy and the listener, speech intelligibility will return. Remove the Sonex and the speech will again be interfered with. By noting the position of the Sonex relative to the listener and the room surfaces, you can usually detect the offending area. Cupping your ears and moving around the room surfaces will also tell much about reflections in the room.

Troy Savings Bank Concert Hall has a 3+ second reverberation time (empty), yet normal conversation level can be heard clearly from the stage at the back of the upper balcony. Troy Concert Hall has no absorption in a conventional sense (all wood) but is very diffuse.

12.6.3 Effect of Reverberation on Intelligibility

The effect of reverberations on intelligibility is far less than single late high level reflections, inadequate *SNR*, or comb filters generated by sources within one foot or less of the primary source.



REFLECTIONS AND COMB FILTERS

$d := 75$ L direct sound path $c := 1130$ L /sec

$h := 30$ L source height $a := 0.5$ dimensionless

Comb filters space linearly therefore nf_n and nf_p calculate the higher frequency notches and peaks.

$d_1 := \sqrt{d^2 + (2\cdot h)^2}$ $d_1 = 96$ L reflected source path

$f_n := \frac{0.5\cdot c}{d_1 - d}$ $f_n = 26.8$ Hz lowest notch frequency

$f_p := \frac{c}{d_1 - d}$ $f_p = 53.7$ Hz first peak frequency

$diff := \frac{0.5\cdot c}{f_n}$ $diff = 21$ L path difference

$d := d_1 - \frac{0.5\cdot c}{f_n}$ $d = 75$ L (used when source distance is unknown but reflected path and f_n are available.)

$d_1 := d + \frac{0.5\cdot c}{f_n}$ $d_1 = 96$ L (used when source distance is known and f_n are available.)

First peak occurs at N λ, first notch at ((2N-1) / 2) λ (Where λ is wavelength)

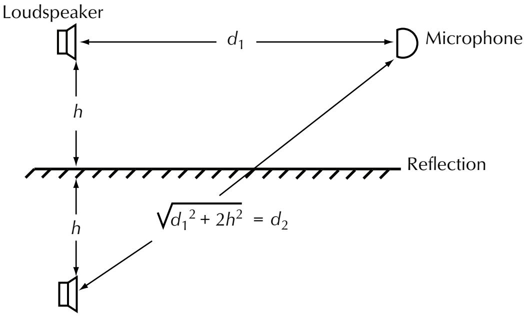
$10\cdot\log\left[\frac{d}{d_1}(1-a)\right] = -4.1$ dB of D/R ratio

Finding distance to nearest reflective surface from lowest notch frequency.

$f_{n1} := 2000$ Hz lowest notch frequency. $c_1 := 13560$ L /sec

$rp := \frac{0.5\cdot c_1}{f_{n1}}$ $rp = 3.39$ L Reflective path distance (rp) Dimension determined by units chosen for c (ie: ft,in,mm,cm,m,etc.)

A. MathCAD program showing reflections and comb filter equations.



B. How to determine path lengths.

Figure 12-6. Reflections and comb filters.

Fig. 12-8 shows three highly audible late reflections (over 100 ms). We were told that this space was so reverberant that it was difficult to use the reinforcement system. Yet, when we put Sonex around the person in this seat and isolated him from the reflections, one could understand clearly the unaided voice from the stage. Without the Sonex, speech was unintelligible.

12.6.4 Variations in the Measurement of Reverberation

TEF analysis has revealed some unexpected details in the measurement of reverberation time. Anyone working with high *Q* transducers (*Q*s of 50+) has subjectively experienced the apparent change in RT_{60}. This has always been explained as due to a

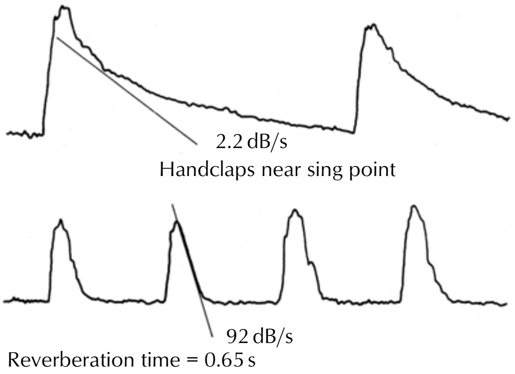


Figure 12-7. Sound system near feedback. (*Courtesy William Snow.*)

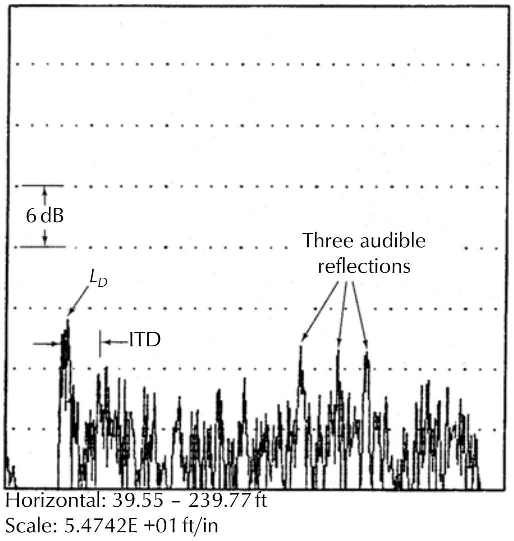


Figure 12-8. Three audible late reflections that seriously affected articulation.

lower reverberant sound field and consequently what decay of energy was present did not last as long before being masked by the ambient noise floor. Various orientations of the sound source reveal that the way a high Q source should be oriented to cover an audience does not excite as many normal room modes as does a lower Q device. Thus it would appear from the evidence available that:

1. The classic method is flawed when being Q dependent.
2. The higher Q sources literally do not excite all the room modes.
3. We possibly need a new, as yet undetermined, method of analyzing the envelope time behavior of large rooms.

Fig. 12-9 shows ETCs of the same room excited from the same source location by three different sources. Source number one has a $Q = 1$. Source number two has a $Q = 5$. Source number three has a $Q = 50$.

Source Number	Q	D_1 (10log Q)
1	1.0	0 dB
2	5.0	7 dB
3	50.0	17 dB

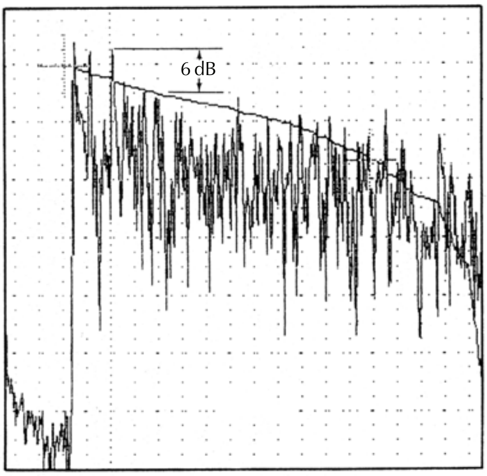
It cannot be overemphasized that analysis reveals that a large number of smaller volume, acoustically absorptive rooms do not develop a reverberant sound field that rises above the ambient noise floor normally present in a space. In such spaces statistical analysis is an exercise in futility. Study of the fine structure of the early reflection, L_{RE}, is of benefit and leads directly to improved performance.

To obtain accurate information about the presence or lack of a reverberant sound field, both sufficient distance from the source must be established and sufficient time must be provided for it to develop, see Fig. 12-10.

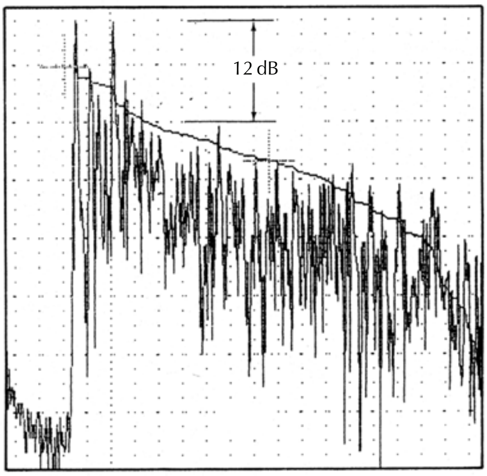
12.7 Critical Distance

One of the most important concepts regarding the statistical acoustic space is D_c. First, we assume that D_c is within 3 dB of the maximum acoustic separation between the microphones in a given room. Again, if we have a microphone in a steady reverberant field, we could wander all over the reverberant area without encountering a sudden change in level that can cause feedback. In fact, we will make use of D_c in determining the following limits in our design of sound systems:

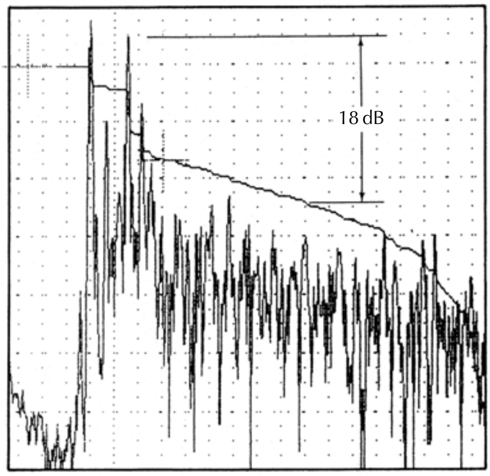
1. The loudspeaker and the microphone should be at least as far apart as D_c. $D_1 = D_c < 45$ ft, where D_1 is the distance between the loudspeaker and the microphone.
2. In rooms with a reverberation time exceeding 1.6 s you will not be able to have any listener beyond 3.16 D_c. As the time raises more, this multiplier will become even lower. (Discussed in detail in Chapter 14 *Designing for Acoustic Gain*. Also see Fig. 12-5C.)



A. Q = 1.

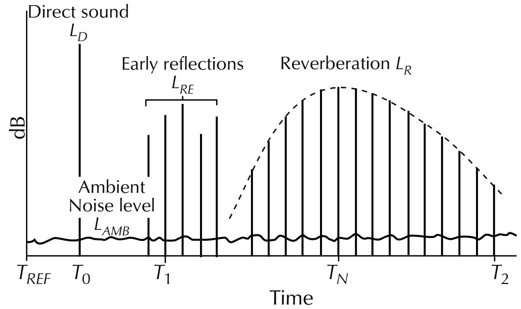


B. Q = 5.
(Note the drop in reverberant level over Q = 1).



C. Q = 50 with a substantial drop in level
of the reverberant field.

Figure 12-9. ETCs made in a room excited from the same source location by three different sources.



$T_0 - T_{REF}$ = Signal travel time to observer (D_0)
$T_1 - T_0$ = Initial Time Delay (ITD) gap
$T_N - T_{REF}$ = Natural room delay
$T_2 - T_1$ = 3-D measurement limits (variable from T_{REF} to T_2)

Figure 12-10. Definition of sound field levels vs. time.

12.7.1 **Q** ***Versus*** Sā ***for Controlling*** **D_c**

In examining the equation for D_c, it is apparent that both *Q* and *Sā* have the same relative weight. This means that in a space that requires a doubling of *Sā* to be acceptable, we could just as well leave *Sā* alone and double *Q*. In typical church systems, for example, the doubling of *Sā* can easily cost $100,000 and change the entire visual appearance of the structure as well as making the music director very unhappy. Doubling *Q* usually costs under $10,000. While the array may be huge, it does occupy only one spot and not whole walls and ceilings. This is a relatively new concept and not widely practiced, though certain acoustic consultants have effectively used the general idea for years. We can now enumerate a few of the factors proceeding from the existence of D_c in a space:

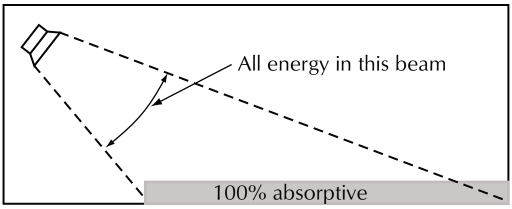
1. D_c determines the maximum acoustic separation hence maximum acoustic gain.
2. D_c determines the ratio of direct-to-reverberant sound.
3. D_c determines the required directivity of the loudspeaker in an already existing room.
4. D_c can determine the required room characteristics in a space being planned if a chosen loudspeaker is desired.

12.7.2 D_c ***Multipliers and Dividers***

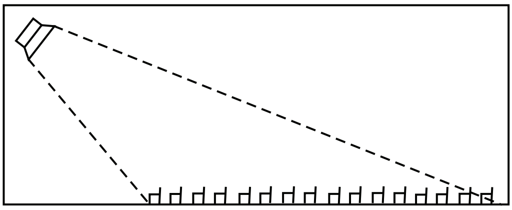
Just as *N* operates as a D_c divider, there are factors that can operate as D_c multipliers. First, let's take an extreme case in which the $C\angle$ contains all the useful energy. It is aimed at an audience area that is 100% absorptive as shown in Fig. 12-11A. For example, if the loudspeaker has an axial *Q* of 5, the room has an

$\bar{a}$ of 0.01, and the audience area has an $\bar{a}$ of 1, the apparent Q would be

$$\begin{aligned} Q_{App} &= Q_{Axial}\left(\frac{1-\bar{a}\text{ of total room}}{1-\bar{a}\text{ of audience area}}\right) \\ &= 5\left(\frac{1-0.01}{1-1}\right) \\ &= \infty \end{aligned} \tag{12-37}$$



A. No reverberation.



B. More typical case.

Figure 12-11. D_c multiplier, *Ma*, illustrated.

Even if the room were highly reverberant, this loudspeaker would cause no reverberation.

A more typical case is shown in Fig. 12-11B. In this case the loudspeaker still has an axial Q of 5, the room has an $\bar{a}$ of 0.16, and the audience area has an $\bar{a}$ of 0.32. The apparent Q would be

$$\begin{aligned} Q_{App} &= Q_{Axial}\left(\frac{1-\bar{a}\text{ of total room}}{1-\bar{a}\text{ of audience area}}\right) \\ &= 5\left(\frac{1-0.16}{1-0.32}\right) \\ &= 6.18 \end{aligned}$$

12.7.3 Theory of Ma

The D_c modifier (*Ma*) results from the removal of additional energy from the signal emitted upon its first encounter with a selected absorbent boundary surface than would have been expected, had the same energy first encountered a surface possessing the average absorption coefficient of the space as a whole.

In the limiting case, if the area the sound energy first encountered were 100% absorptive and if none of the energy encountered any other surface, there would be no reverberation. Thus, such an *Ma* makes the source act as if it were in a free field. It would appear that *Ma* is only of interest for some very high Q devices.

Consider what is actually happening at a listener's ears. As *Ma* increases the ratio of direct-to-reverberant sound heard increases if the listener is situated on the absorbing surface that the energy first encounters. The main purpose of increasing Q or *Ma* is to increase this ratio at the listeners ears.

Another parameter available to us that allows us to accomplish the same results is to move the loudspeaker closer to the listener. If we move a loudspeaker of any given Q in a room of any given *Ma* to half its former distance from the listener's ears, we raise the direct-to-reverberant ratio by 6 dB. This could also be accomplished by leaving the loudspeaker at its original position and raising its Q by a factor of four.

The relation of the loudspeaker's angular discrimination relative to a microphone's angular discrimination is referred to as an electroacoustic modifier of D_c (*Me*) and is discussed in detail in Chapter 18 *Loudspeakers and Loudspeaker Arrays* which discusses the planning of loudspeaker arrays and acoustic gain.

Additionally, a tilted rear wall may raise the apparent Q at the rear seats in an auditorium even more. When you measure Q in a room with a calibrated loudspeaker, the difference between what you calculate and the calibration includes all the multipliers and dividers. Therefore, such a calibrated test source allows you to do several things:

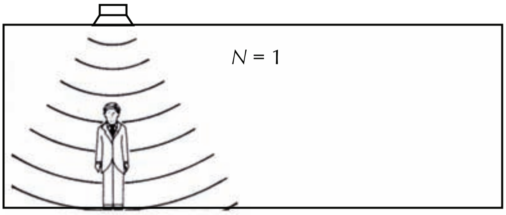
1. Measure the room absorption, $S\bar{a}$, by the reverberation time method and calculation. Then measure the apparent Q. This will include all multipliers and dividers. By changing such variables as position of materials and position of sources, you can investigate the effect of such phenomena.
2. Measure the room absorption, $S\bar{a}$, by the critical-distance method using the Q of your calibrated sources in your calculations. By using this room absorption, which includes all the multipliers and dividers, you can accurately measure the axial Q of unknown loudspeakers.

12.7.4 The Effect of the N Factor

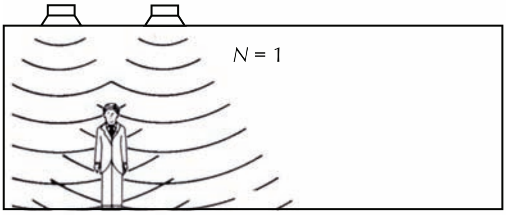
N is the ratio of the acoustic power produced by all loudspeakers to the acoustic power produced by the loudspeaker or loudspeaker groups providing the listener with direct sounds. In its simplest form as shown in Fig. 12-12, two loudspeakers are furnishing direct sound to the listener and there are four loudspeakers producing equal acoustic power. Therefore

$$N = \frac{\text{Total number of loudspeakers}}{\text{Number of loudspeakers producing direct sound to the listener}}$$

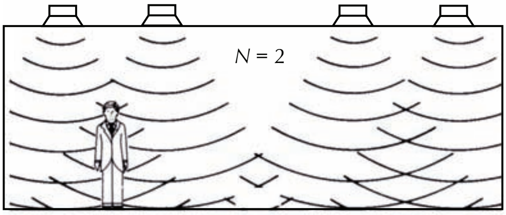
$$= 2$$



Direct to reverberant ratio = 17 dB



Direct to reverberant ratio = 17 dB



Direct to reverberant ratio = 14 dB

Figure 12-12. Visualization of *N*.

To demonstrate the “*N*” effect, sound fields were compared using a loudspeaker on stage as a substitute for a talker into the sound system microphone, Figs. 12-13,12-14 and 12-15.

1. Loudspeaker on stage only (a talker from the stage).
2. Loudspeaker on stage plus a center cluster ($Q = 11$).
3. Loudspeaker on stage plus two low-Q stereo loudspeakers ($Q = 2$).
4. All loudspeakers on at the same time.

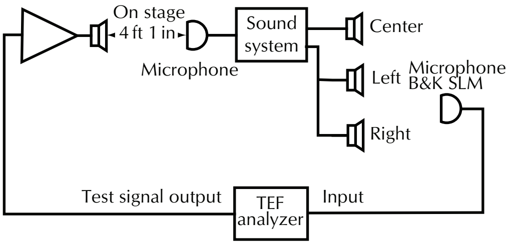
The effects of the complex *N* factor are clearly evident. Some interesting effects from discrete early reflections that affect the ratio of direct (defined here as the first 50 ms) sound level versus reverberant sound level are also apparent. The effect of *N* is a 10log function. If all four sources developed equal acoustic power then the expected deterioration would be 10log 5 ≈ 7 dB.

This rule follows within a few decibels at all sampled locations, see Table 12-1.

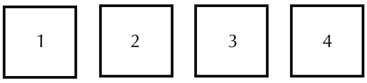
Table 12-1. Summary of Direct to Reverberant Ratios

Row	Stage	Center	Stereo	All
1	14	6	4	4
6	14	7	3	3
11	13	5	5	4
16	12	4	5	3
21	10	9	0	4
26	9	6	4	3
31	8	4	1	1

Direct = 0–50 ms
Reverberant = 50 ms–444 ms



Test setup.



1. Loudspeaker on stage only
2. 1 + center cluster
3. 1 + L & R stereo
4. 1 + 2 + 3

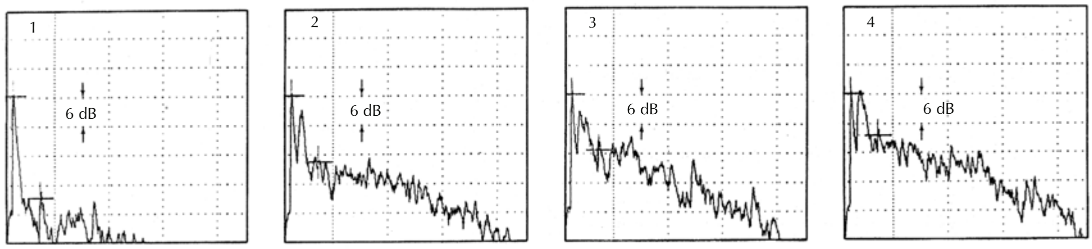
Frequency range: 1500 Hz–2400 Hz
Time span: 444 ms with 10 ms offset (10 ms–454 ms)
Sweep rate: 200 Hz
Single pole integration: 5 ms
Microphone locations: Rows 1, 6, 11, 16, 21, 31 (last)
Center: 2 Renkus Heinz GBH1250-9 83° x 50°; Q=11
L & R: Bes Q = 2

Measurement parameters.

Figure 12-13. Measurement parameters for sound system study in Fig. 12-14

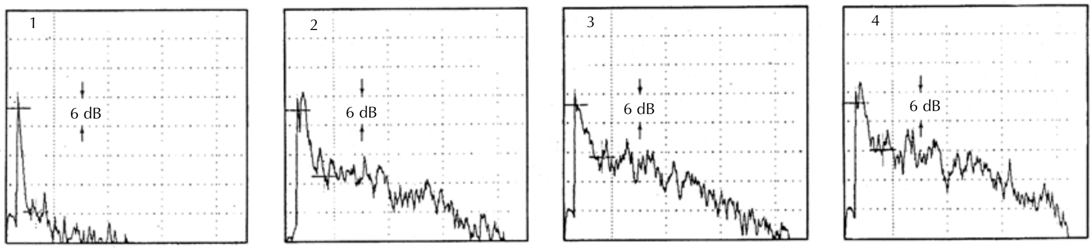
12.7.5 Factors to Watch for in Rooms

The following factors can be serious trouble for the sound system if they are not properly controlled:



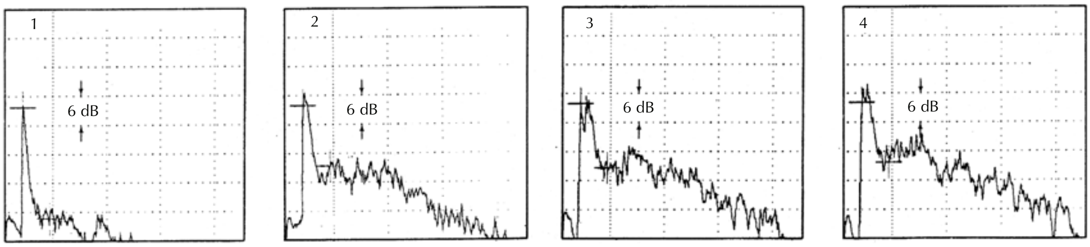
Direct to reverberant ratio: 14.32 dB
Direct to reverberant ratio: 5.95 dB
Direct to reverberant ratio: 4.13 dB
Direct to reverberant ratio: 4.08 dB

Row 1



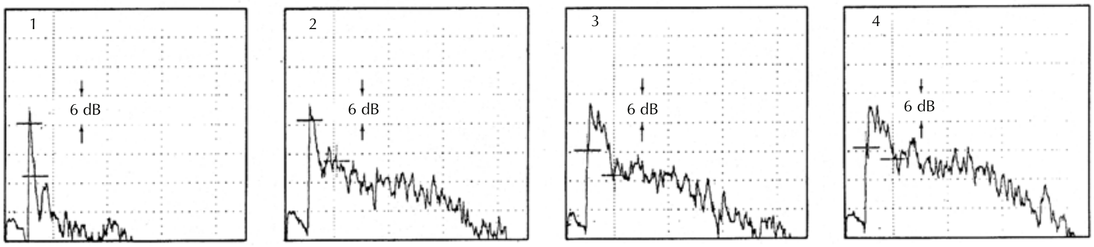
Direct to reverberant ratio: 14.40 dB
Direct to reverberant ratio: 7.15 dB
Direct to reverberant ratio: 3.39 dB
Direct to reverberant ratio: 3.49 dB

Row 6



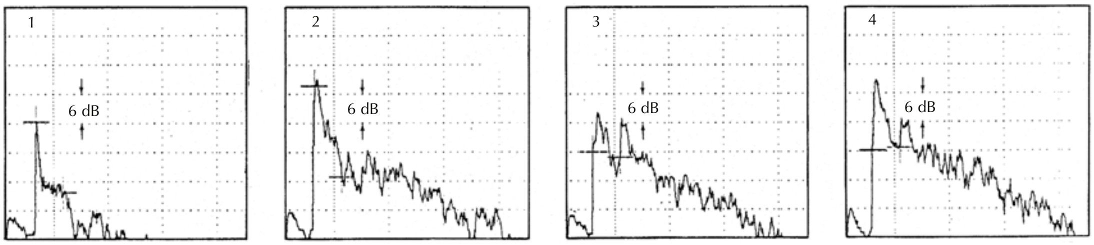
Direct to reverberant ratio: 13.31 dB
Direct to reverberant ratio: 5.49 dB
Direct to reverberant ratio: 4.74 dB
Direct to reverberant ratio: 4.06 dB

Row 11



Direct to reverberant ratio: 12.25 dB
Direct to reverberant ratio: 4.03 dB
Direct to reverberant ratio: 4.96 dB
Direct to reverberant ratio: 2.73 dB

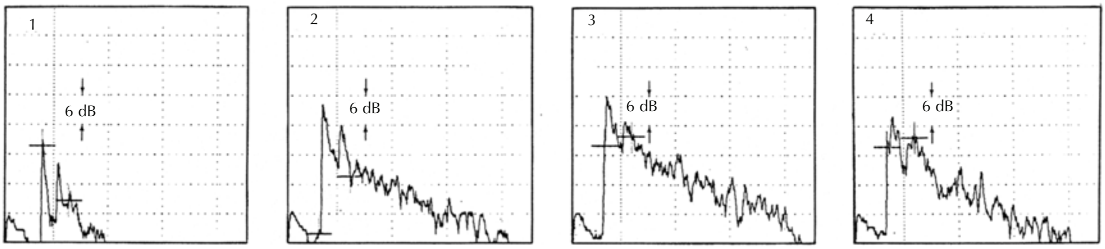
Row 16



Direct to reverberant ratio: 10.24 dB
Direct to reverberant ratio: 8.83 dB
Direct to reverberant ratio: 0.03 dB
Direct to reverberant ratio: 4.31 dB

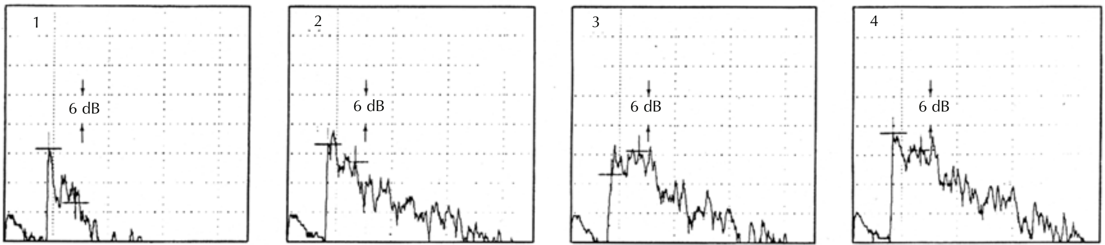
Row 21

Figure 12-14. Complete study of a sound system where the direct-to-reverberant ratio is degraded as each loudspeaker is added. Prepared by Rollins Brook of BBN.



Direct to reverberant ratio: 8.55 dB Direct to reverberant ratio: 6.17 dB Direct to reverberant ratio: 4.17 dB Direct to reverberant ratio: 3.31 dB

Row 26



Direct to reverberant ratio: 7.51 dB Direct to reverberant ratio: 4.29 dB Direct to reverberant ratio: 0.87 dB Direct to reverberant ratio: 1.17 dB

Row 31

Figure 12-14. (cont.) Complete study of a sound system where the direct-to-reverberant ratio is degraded as each loudspeaker is added. Prepared by Rollins Brook of BBN.

1. Curved surfaces, especially concave curved surfaces.
2. Absolutely parallel walls. Such walls cause flutter-echo. A splay of 1 inch per foot will avoid this problem.
3. Absorption on the ceiling. Unless the ceiling is very high (over 60 ft), the placement of absorption on it means the sound system has lost some useful reflecting surfaces. Absorption belongs on rear walls (large spaces only), rear ceilings, in the seats, etc.
4. Potential ambient noise sources—air handlers, unenclosed machinery, etc.
5. Extra wide or round audience seating.

12.8 Conclusion

The meaningful use of acoustical absorption is not limited to its statistical application. Indeed sound systems are more often installed in spaces where the statistical equations are invalid than in spaces where they are valid.

In semireverberant spaces and in very "dead" spaces the use of absorption to control discrete specular reflections is quite valid in spite of the application of the material having no real meaning in a statistical sense. Even in spaces where the statistical equations are valid, intelligibility can be, and often is, degraded by a specular reflection which must, of course, be isolated and corrected directly, not statistically. Therefore, as we proceed into Chapter 13 *Small Room Acoustics*, it is well to bear in mind that many "large rooms" have some small room properties at certain frequencies, especially with regard to specular reflections.

Bibliography

D. Davis. "Contemporary Electro-Acoustic Investigations," *Syn-Aud-Con Tech Topics*, Vol. 7, No. 11 (1980).

_______. "Uses, Abuses, and Misuses of the Critical Distance Concept," *Syn-Aud-Con Tech Topics*, Vol. 7, No. 12, (1980).

D. Davis and C. Davis, "What Reverberation is and What it is Not," *Syn-Aud-Con Tech Topics*, Vol. 12, No. 13 (1985).

_______. "Sabine's Reverberation Time and Ergodic Auditorium," *J. Acousti. Soc. Am.*, Vol. 58 (1975), pp. 643-655.

W. B. Joyce. “Power Series for the Reverberation Time.” Paper presented at 97th Convention of the Acoustical Society of America, Cambridge, MA (June 13, 1979).

D. L. Klepper. “Sound Systems in Reverberant Rooms for Worship,” *J. Audio Eng. Soc.*, Vol. 18 (Aug. 1970).

_______. “Improved Reverberation Time Calculations,” *Syn-Aud-Con Tech Topics*, Vol. 6, No. 14 (1979).

R. F. Norris. “Appendix II: A Discussion of the True Coefficient of Sound Absorption—A Derivation of the Reverberation Formula,” *Architectural Acoustics*, pp. 603-665. New York: Wiley, 1932.

B. Rayleigh. *The Theory of Sound*, Vols. I and II, 2nd ed. New York: Dover, 1945.

W. C. Sabine. Collected Papers on Acoustics, Cambridge, MA: Harvard Univ. Press, 1922.

M. R. Schroeder. “Computers in Acoustics: Symbiosis of an Old Science and a New Tool,” *J. Acoust. Soc. Am.*, Vol. 45, No. 5 (1969).

R. S. Shankland. “Acoustics of Greek Theaters,” *Physics Today* (Oct. 1973), pp. 30-35.

William B. Snow. “Frequency Characteristics of a Sound Reinforcing System,” *J. Acoust. Soc. Am.* (April 1955).

Chapter 13

Small Room Acoustics

by Don Davis

13.1 Non-Statistical Spaces . . . 237
13.2 Small Room Acoustical Parameters . . . 238
13.3 Small Room Reverberation Times . . . 238
13.4 Small Room Resonances . . . 239
13.5 Modes . . . 239
  Damped and Undamped Modes . . . 239
  Modal Decay Rates . . . 239
13.6 What Is an Eigen Mode? . . . 239
  Normal Modes Defined . . . 240
13.7 Small Room Geometry . . . 240
  Desirable Room Ratios . . . 240
13.8 The Initial Signal Delay Gap (*ISD*) . . . 240
  Selecting an Initial Signal Delay Gap (*ISD*) . . . 241
13.9 Reflections . . . 242
  Useful Definitions . . . 242
13.10 Reflection Free Zone . . . 244
  Development of the Reflection Zone . . . 244
13.11 Diffusion . . . 245
  Quadratic Residue Diffusors . . . 248
13.12 Conclusion . . . 248
Bibliography . . . 248

13.1 Non-Statistical Spaces

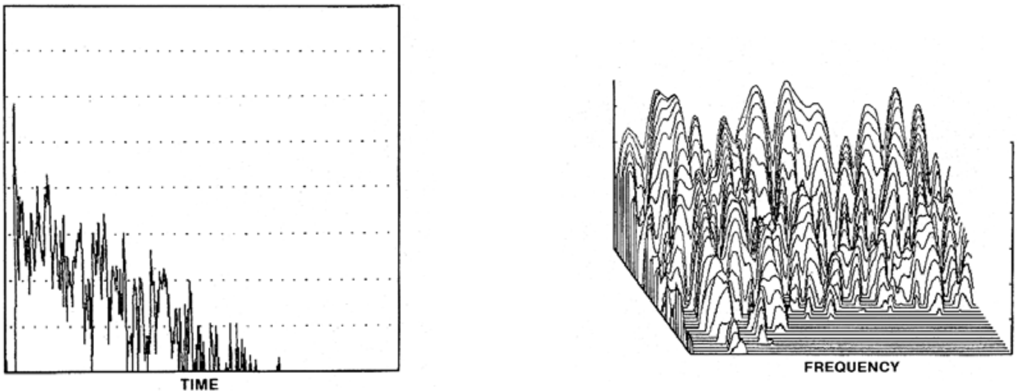
Sound systems are more frequently installed in spaces where the statistical equations are invalid than in spaces where they are valid. Consequently, the meaningful use of acoustic absorption is not limited to its statistical application. We use absorption to control discrete specular reflections in semi-reverberant and very dead spaces in spite of the fact that the material has no statistical meaning. Even in spaces where the statistical equations are valid, intelligibility can be degraded by a specular reflection that must be isolated and corrected directly, not statistically. Therefore, as we examine the properties of small-room acoustics bear in mind that many large rooms have small room properties at certain frequencies, especially with regard to specular reflections. The "acoustical" size of a room is a frequency dependent phenomenon.

We are dealing here with room modes rather than a statistical reverberant sound field. One glance at the illustrations of the 3-D TEF plots of the reflected sound in an acoustically "small" room and a "large" room reveals the dramatic differences in the energy density with time, see Fig. 13-1.

How many reflections can occur in such a small space in 0.1 s?

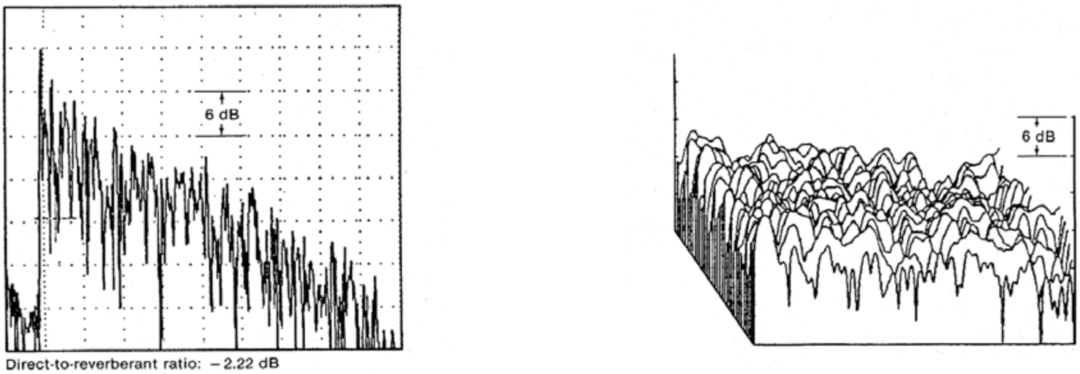
$$\begin{aligned} MFP &= \frac{4V}{S} \\ &= \frac{4(2288\ \text{ft}^2)}{672\ \text{ft}^2} \qquad (13\text{-}1) \\ &= 16\ \text{ft} \end{aligned}$$

In 0.1 s the sound travels 0.1 s (1130 ft/s) = 113 ft so the number of reflections is



A. ETC of a small room showing lack of a dense field of reflections. After 150,000 µs 150 ms or 150 ft, all reflections have died out. (*Courtesy Charles Bilello.*)

B. 3-D of same room without a reverberant sound field but with room modes. (*Courtesy Charles Bilello.*)



C. ETC of a large room where a sound field is still present at 1,134 ms or approximately 1000 ft. (*Courtesy Ruth Eckerd Hall.*)

D. 3-D of a large concert hall with a good reverberant field.

Figure 13-1. Proof that there is a fundamental difference between small and large reverberant spaces.

$$\text{No of reflections} = \frac{\text{Distance the sound travels}}{\text{MFP}} = \frac{113\text{ ft}}{16\text{ ft}} = 7$$

Obviously this does not comprise a mixing, homogeneous, statistical reverberant sound field and indeed the illustrations demonstrate this quite effectively. f_c is the acoustic juncture between large and small rooms. f_c coincides with the dimension of the room equal to the lowest wavelength that can fully develop across that dimension. In physically small rooms, f_c can be as high as 500 Hz, whereas it falls below 30 Hz in physically large rooms. Bolt, Beranek and Newman developed an important chart that it calls controllers of steady-state room acoustic response, Fig. 13-2.

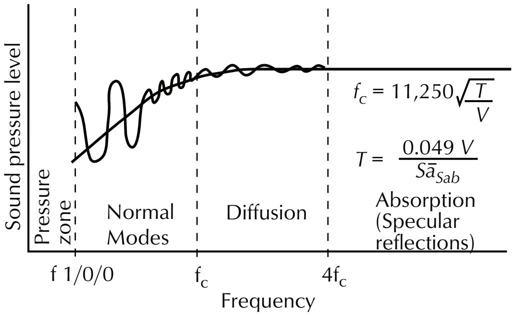


Figure 13-2. Controllers of steady-state room acoustic response. (*Courtesy Bolt, Beranek, and Newman.*)

The frequency dependency of the pressure zone, the modal zone, the diffusion zone, and the specular reflection zone determine, how room treatment is used. In rooms that are both physically and acoustically small, the pressure zone may be useful to nearly 100 Hz. Diaphragmatic absorbers are useful from 80 Hz to perhaps 500 Hz, whereas quadratic residue diffusors are useful from 500 Hz to 2000 Hz. Above 2000 Hz discrete reflections must be specifically controlled. Application of a good technique at an incorrect frequency is as disastrous as choosing the wrong technique.

13.2 Small Room Acoustical Parameters

It is rare to need to reinforce in these small spaces as even a weak voice can carry 12 to 20 ft. Teleconferencing may change this but primarily even in small meeting rooms with teleconferencing "soft switching" is employed, making the system into essentially a reproduction rather than a reinforcement system. Because of this, looking at these small rooms in terms of their use by live talkers ($Q = 2.5$ in the articulation frequency region) is a good starting point for designers. Free field equations plus identification and tracking of early reflections are usually the total environmental analysis required in terms of the usable sound field. Ambient noise level, room geometry, and any unusual modifiers (i.e., a totally absorptive rear wall) receive the usual consideration.

13.3 Small Room Reverberation Times

To quote the late Ted Schultz (formerly of BB&N):

> *In a large room, if one has a sound source whose power output is known, one can determine the total amount of absorption in the room by measuring the average pressure throughout the room. This total absorption can then be used to calculate the reverberation time from the Sabine formula. This method fails badly in a small room, however, where a large part of the spectrum of interest lies in a frequency range where the resonant modes of the room do not overlap but may be isolated.... In this case the microphone, instead of responding to a random sound field (as required for the validity of the theory on which these methods depend), will delineate a transfer function of the room.... It does not provide a valid measurement of the reverberation time in the room.*[*]

What is often overlooked in the attempted measurement of RT_{60} in small rooms is that the definition of RT_{60} has two parts, the first of which is unfortunately commonly overlooked.

1. RT_{60} is the measurement of the decay time of a well-mixed reverberant sound field well beyond D_c.
2. RT_{60} is the time in seconds for the reverberant sound field to decay 60 dB after the sound source is shut off.

Since, in small rooms, there is no D_c, no well-mixed sound field, hence, no reverberation but merely a series of early reflected energy, the measurement of RT_{60} becomes meaningless in such environments.

---

[*] T. Schultz, ASA unpublished paper, 1984.

What becomes most meaningful is the control of the early reflections because there is no reverberation to mask them.

13.4 Small Room Resonances

Many of us have listened in small rooms to the low frequency resonances that occur when one of the dimensions of the room supports a particular frequency like a "tuned" tube. Fig. 13-3A is such a resonance (about 125 Hz). Fig. 13-3B is the same measurement made after the construction of a diaphragmatic absorber.

13.5 Modes

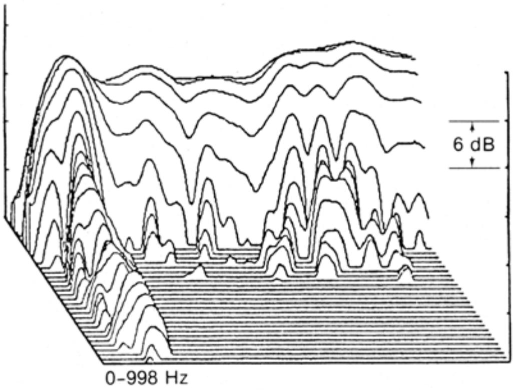
13.5.1 Damped and Undamped Modes

In Fig. 13-4, the effect of "undamped" modes are plotted, as a decay time, for a small broadcast studio. The damping is provided by diaphragmatic absorption at the lower frequencies. Such low frequency absorption (the flexing of panels at the low frequencies passing the energy from the small room to a calculated cavity) not only reduces the peak amplitude of the mode but broadens its bandwidth (lowers its resonant Q).

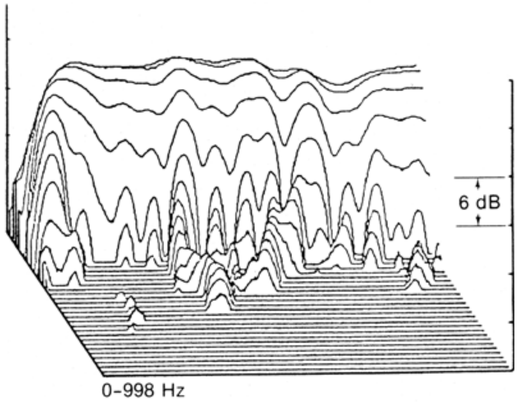
13.5.2 Modal Decay Rates

Fundamental point: modal decay rates are not reverberation. Reverberation is "the time in seconds that it takes a diffuse sound field, well beyond a real critical distance, to lower in level by 60 dB when the sound source is turned off." Modal decay rates are dB-per-second (dB/s) rate of decay for a specific modal frequency.

Eigen Modes are sometimes referred to as "Eigen Tones" which has led some users in the United States to regard them as "Eigen Frequencies." This is a dangerous misconception inasmuch as they are dependent upon wavelength and as the velocity of sound varies as the temperature in the space, Eigen Modes shift in apparent frequency in order to maintain the same wavelength relationship with the boundary surfaces. A more correct English translation would be "Eigen Wavelengths."



A. Resonance at 125 Hz



B. After construction of a Helmholtz resonator

Figure 13-3. Small broadcast control room. (*Courtesy Doug Jones and WFMT.)*

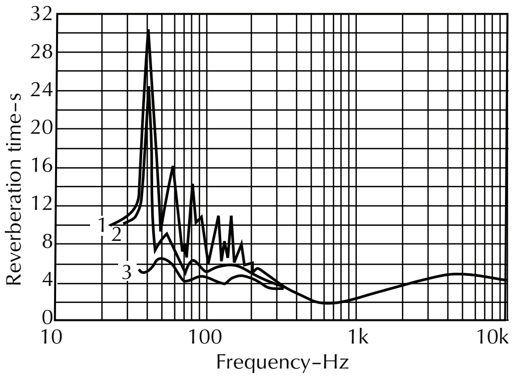


Figure 13-4. Controlling normal mode damping and bandwidth.

13.6 What Is an Eigen Mode?

An Eigen Mode is the European name for a standing wave. Standing waves are dependent upon the internal dimensions of an enclosure. The first mode will be found at

$$f_0 = \frac{c}{2L} = \frac{c}{\lambda} \tag{13-2}$$

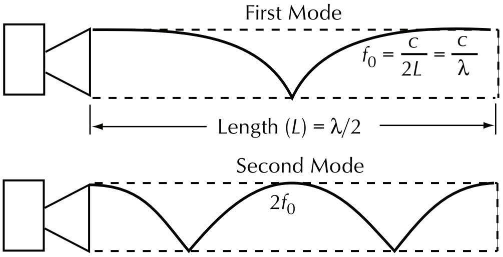
where,

f_o is the frequency in Hz of the first mode,

λ is the wavelength of the frequency and is equal to twice the longest dimension of the enclosure,

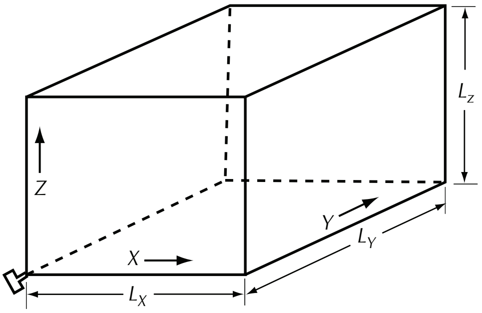
c is the velocity of sound in the air.

See Figs. 13-5 through 13-7 for standing waves.



Third Mode = 3/0, Fourth Mode = 4/0, etc.

Figure 13-5. Generating standing waves.



Tiny source of sound at x = y = z = 0
where N_x, N_y, and N_z can be 0, 1, 2, 3, ...

$$f_N = \frac{c}{2}\sqrt{\left(\frac{N_x}{L_x}\right)^2 + \left(\frac{N_y}{L_y}\right)^2 + \left(\frac{N_z}{L_z}\right)^2}$$

Figure 13-6. Normal modes in a rectangular room where the Eigen Tones are generated by standing waves.

13.6.1 Normal Modes Defined

Axial modes. in which the component waves move parallel to an axis (one dimensional), the (N_x, 0, 0), (0, N_y, 0), and (0, 0, N_z) modes of vibration.

Tangential modes. in which the component waves are tangential to one pair of surfaces, but are oblique to the other two pairs (two dimensional), the (N_x, N_y, 0), (N_x, 0, N_z) and (0, N_y, N_z) modes of vibration.

Oblique modes. in which the component waves are oblique to all three pairs of the walls (three dimensional), the (N_x, N_y, N_z) modes of vibration.

Plot axial, tangential, and oblique modes independently on three lines, Fig. 13-8.

13.7 Small Room Geometry

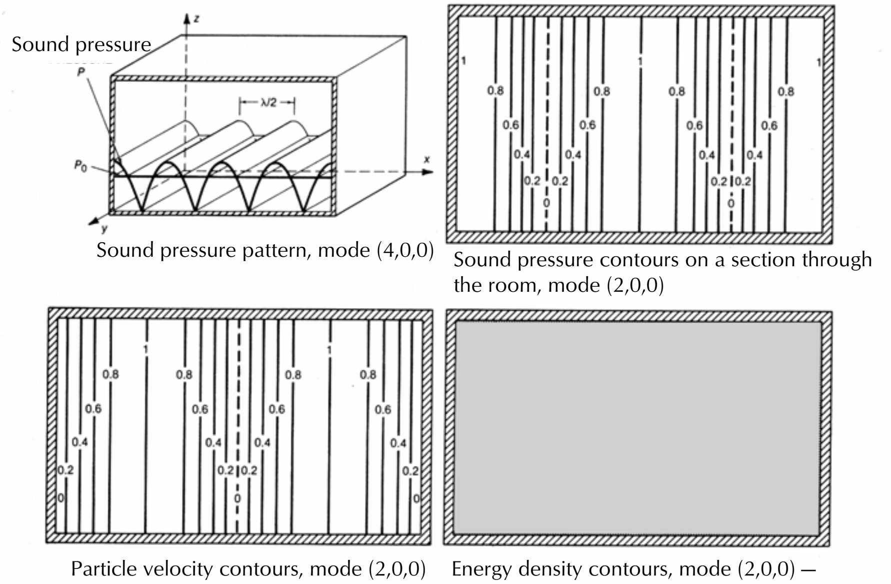
13.7.1 Desirable Room Ratios

This might more meaningfully read "Undesirable Room Ratios." Figs. 13-9 and 13-10 show, within the enclosed curve, the ratios that have been found to be acceptable. Here the criterion is simply to avoid falling outside the enclosed area in your basic room dimensional ratios.

13.8 The Initial Signal Delay Gap (ISD)

The initial signal delay gap is a fundamental room parameter. It was first clearly identified and described by Leo J. Beranek. This "gap" is defined as the time between the arrival of the direct sound, L_D, at a listener's ears and the arrival of the first significant reflection. "Significant" is intended to mean the first reflection whose level approximates that of the peak of the exponentially growing and decaying reverberant sound field. Since in a small room we don't have a reverberant sound field as defined in the classic sense, we look for the first reflection within 6 dB of the highest level reflection.

In small rooms, the ISD is normally quite short, on the order of 1 to 5 ms. In a special design of control rooms for monitoring recording studios, a principle called "Live End Dead End," LEDE, is used that allows ISDs of from 10 to 20 ms to be developed in rooms with dimensions as small as 2000 ft^3. Before the use of the TEF analyzer during the building of LEDE rooms, the front half of the room was made as absorptive as possible (only the diffusion and spectral frequencies are necessary) and the other half (the half to the rear of the listener) was (and still is) made as reflective and diffusive (highly important and often overlooked) as possible. It's in the reflective half that a thorough plotting of reflections becomes a necessity. See Table 13-1 for



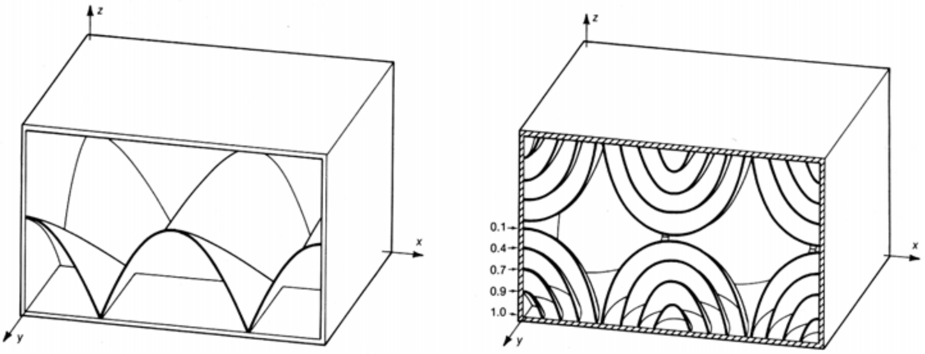
Sound pressure pattern, mode (4,0,0)

Sound pressure contours on a section through the room, mode (2,0,0)

Particle velocity contours, mode (2,0,0)

Energy density contours, mode (2,0,0) — For an axial mode the energy density is uniform.

A. Axial mode. (*Courtesy B & K Technical Review.*)



B. Tangential mode

C. Oblique mode

Figure 13-7. Sound distribution in a rectangular room.

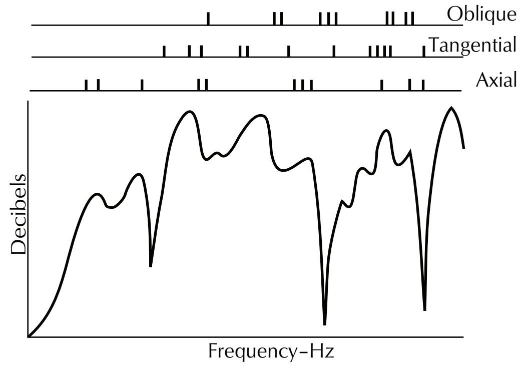


Figure 13-8. Plotting modal wavelengths as frequencies.

sound fields present in a control room and combining of closely spaced signals.

13.8.1 Selecting an Initial Signal Delay Gap (ISD)

To select an initial signal delay (ISD) gap you need to know the ISD of the studio or other environment surrounding the musicians.

Additionally, the control room's ISD must be made longer than the studio's ISD if it is to allow reproduction of the studio's gap over the monitor loudspeakers, Figs. 13-11 and 13-12.

The control room's first significant reflection should fall within the Haas zone (see Chapter 15 *Designing for Intelligibility* for a discussion of the Haas zone). Experience indicates that all subsequent reflected energy should appear as a sloping straight line on the analyzer's display of the

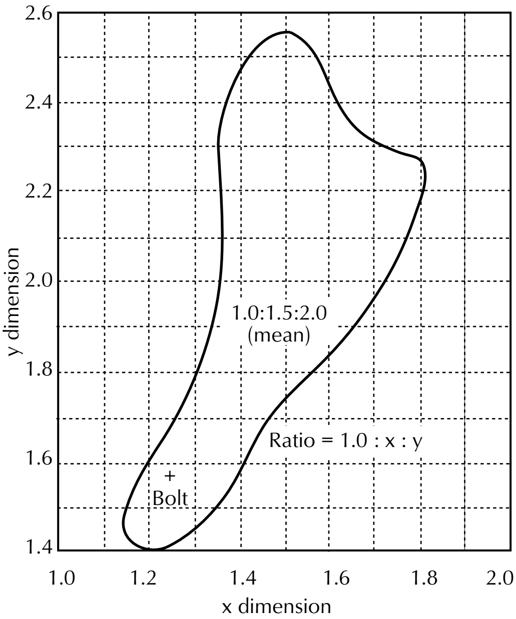


Figure 13-9. Acceptable room ratios. Courtesy Bolt, Beranek, and Newman.

ASHRAE:	1 : 1.17 : 1.47
	1 : 1.45 : 2.10
BOLT:	1 : 1.28 : 1.54
IAC:	1 : 1.25 : 1.60
SEPMEYER	1 : 1.14 : 1.39
$1 : \sqrt[3]{2} : \sqrt[2]{2}$	1 : 1.26 : 1.41

Figure 13-10. Recommended small room dimension ratios.

Envelope Time Curve, ETC, i.e., an exponential decay rate. It is also known that energy that exceeds this slope is detrimental and audible, particularly so should its time interval fall outside the "Haas zone," Fig. 13-13. The difference between the ETC of a small room (control room) and a concert hall is illustrated in Fig. 13-14. Note the time scales and the difference in ISD gaps.

Table 13-1. Sound Fields Present in Control Rooms

$$L_T = 10\log(10^{L_D/10} + 10^{L_R/10} + 10^{L_n/10}) \tag{13-3}$$

$$L_R = 10\log(10^{L_T/10} - 10^{L_n/10} - 10^{L_D/10}) \tag{13-4}$$

where,
L_T is the total sound in decibels,
L_D is the direct sound level in decibels,
L_R is the reverberant sound level in decibels,
L_n is the ambient noise level in decibels.
All levels are in decibels re 20 μPa.
The addition of two signals having the same frequency but different levels and phases is given by:

$$L_{comb} = 20\log\sqrt{\left(10^{\frac{L_1}{20}}\right)^2 + \left(10^{\frac{L_2}{20}}\right)^2 + 2\left[\left(10^{\frac{L_1}{20}}\right)\left(10^{\frac{L_2}{20}}\right)(\cos(a_1 - a_2))\right]} \tag{13-5}$$

where,
L_{comb} = combined sound level of two signals in decibels,
L_1 = sound level of first signal in decibels,
L_2 = sound level of second signal in decibels,
a_1 = phase angle of L_1,
a_2 = phase angle of L_2.

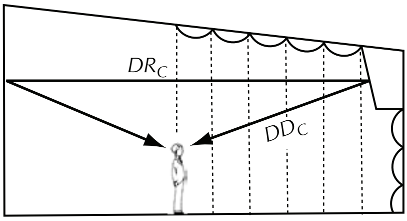
The following illustrates the effect of phase on equal-level signals. A 6 dB addition signifies a dominant direct sound field at a precise point. It does not signify a sound power increase over a given area.

L_1	a_1 (deg)	L_2	a_2 (deg)	L_{comb} (dB)
90	0	90	0	96.02
90	0	90	10	95.99
90	0	90	20	95.89
90	0	90	30	95.72
90	0	90	40	95.48
90	0	90	50	95.17
90	0	90	60	94.77
90	0	90	70	94.29
90	0	90	80	93.71
90	0	90	90	93.01
90	0	90	100	92.18
90	0	90	110	91.19
90	0	90	120	90.00
90	0	90	130	88.54
90	0	90	140	86.70
90	0	90	150	84.28
90	0	90	160	80.81
90	0	90	170	74.83
90	0	90	180	$-\infty$

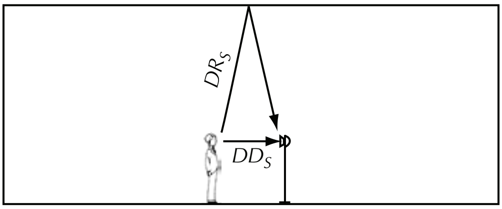
13.9 Reflections

13.9.1 Useful Definitions

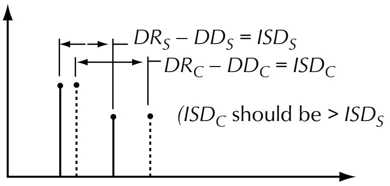
The LEDE concepts are physically simple but psychoacoustically complex. The goal of an LEDE control room is to let mixing engineers, who sit at the console, hear the first reflections from the recording studio over the control room loudspeakers before they hear any from the control room they are sitting in.



A. LEDE control room.



B. Studio.



DR_S = Distance (or time) first reflection travels in studio.
DR_C = Distance (or time) first reflection travels in control room.
DD_S = Distance (or time) direct sound travels in studio.
DD_C = Distance (or time) direct sound travels in control.

Figure 13-11. ISD in a studio and a control room.

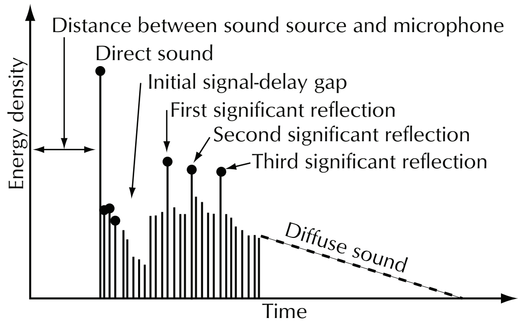


Figure 13-12. Energy density versus time for an LEDE control room.

Dead End. This simply means that no early reflections can be allowed to occur in the front half of the control room. This may be accomplished by using absorption, reflection-free zones (RFZ), or any other method that meets the criteria of having no reflections from the front half of the control room before the deliberately installed diffuse energy occurs.

Live End. The live end consists of three functions:

a. ***Haas effect***—A first reflection strong enough from the studio, as heard over the control room's loudspeakers, to provide the Haas effect.

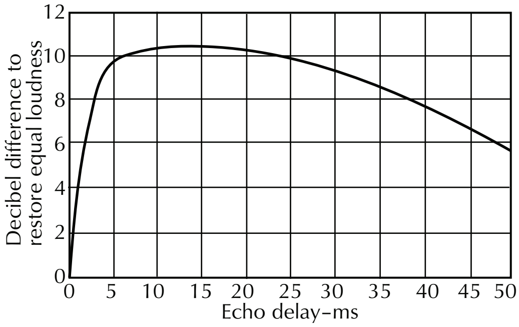
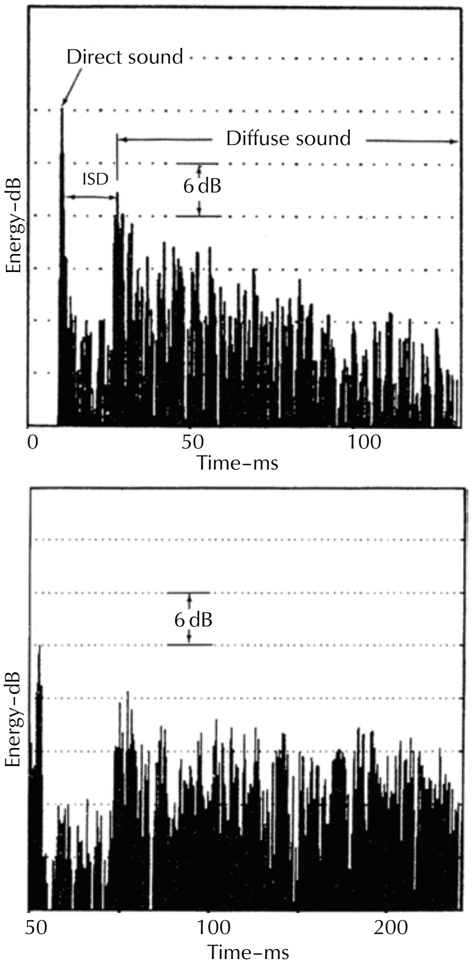


Figure 13-13. Häas zone for a single reflection.



ETCs of Master Sound Astoria recording studio (top) and a great concert hall (bottom). Note the clean ISD level of the clump of early reflections compared to the direct sound. Also note the exponential decay of ETCs of Master Sound Astoria recording studio (top) and a great concert hall (bottom). Note the clean ISD level of the clump of early reflections compared to the direct sound. Also note the exponential decay of the diffused sound.

Figure 13-14. Master Sound Astoria recording studio. (*Courtesy Charles Bilello.*)

b. ***Diffusion***—Our preference is for Schroeder's quadratic residue diffusors (QRD). We believe that the optimum placement of these is behind the mixer's position at the rear wall—a good distance from the mixer is from 7 to 15 ft. The diffusors *should not* be in the on-axis path of the monitors since undesired specular reflections at frequencies above the diffusion frequencies can occur.

c. ***Specular reflectors***—Care must be taken to provide subsequent early reflections. Each reflection should follow an exponential change in level with increasing time. Be sure not to space them at equal delay intervals. These specular reflections should again drive signals back into the diffusors so that the entire audible decay period is diffuse.

13.10 Reflection Free Zone

When studio designers began to use analyzers in the control room during the building and during retrofitting of the control room, it was found that less absorption was needed in the front of the control room to control early reflections.

It is necessary to create a reflection free zone, as Peter D'Antonio calls it, Fig. 13-15. When early control room designers were not using TEF, it was necessary to make the entire front half absorptive.

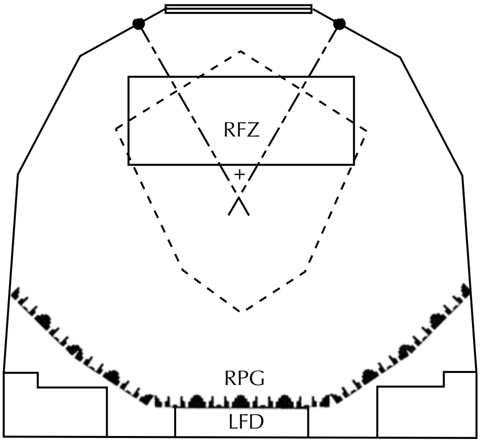


Figure 13-15. Plan view of an RFZ/RPG control room with low-frequency diffusers. Limiting reflections from surface boundaries form a symmetrical six-sided RFZ. (*Courtesy Peter D'Antonio.*)

Comb filters caused by early reflections mimic the pinnae and torso transfer function, which is very destructive to stereo imagery. If the room can be constructed to keep sound off the walls and ceiling, it is not necessary to use heavy absorption in the front of the room. If speakers do cause early reflections, absorption can be strategically placed to minimize the early reflections. Reflections less than 1 ms occur from reflections off the console, from the face of speaker cabinets, and from near-field monitors. Absorption is the best answer.

13.10.1 Development of the Reflection Zone

It has been a synergetic development. A little history would help.

Carolyn "Puddie" Rodgers received her doctorate from Northwestern University in localization and pinna transformations. At Northwestern she worked with Gary Kendall.

Shortly after receiving her Ph.D. she attended a Heyser TDS class. When she saw the comb filters generated by misaligned loudspeakers* on the TDS she remarked that it gave her an idea in future research. To quote from the abstract of her *AES Journal* article (April 1981, Vol. 29, No. 4, p. 226–234)

> *Many studies have shown that the pinna transform incoming signals, superimposing upon the original signal a comb-filter-like spectrum. This spectral shaping has been shown to add an additional cue to the now classic hierarchy of localization cues: interaural intensity, phase, and time of arrival differences. Recent evaluations of misaligned* loudspeakers using time delay spectrometry reveal spectral shapes which are strikingly similar to pinna transformations. The implication is that misaligned loudspeakers*, poorly placed microphones, or other early reflections introduce spectral aberrations which may be decoded by the auditory system as cues to source position. The possible consequences of the pinna transformations to the interpretation of psychoacoustic phenomena such as auditory imaging, the cocktail party effect, and the precedence effect are discussed.*

Doug Jones of Chicago, working with a TEF analyzer in the retrofit of control rooms, met Gary Kendall at Northwestern. A fine paper was written by Doug Jones, Gary Kendall, and William Martens. They performed the following work.

---

*The term misaligned as used here refers to drivers whose acoustic origins are at different distances from the listener.

The design of controlled listening environments is an often overlooked part of the total environment in which computer musicians work. Recording engineers know the importance of being able to listen to their production and mixing work in an environment that supports clarity of image and which sounds the same every time they use it. At the Northwestern University Computer Music Studio we have recently finished construction of a sound room that we intend to use both as a general audio listening environment and as a controlled sound environment for psychoacoustic research into localization.

Experience in many studio monitoring rooms has established that good audio imaging requires the control of reflections. This knowledge has been put to use in the contemporary style of control room design called “Live End—Dead End” or “LEDE.” Unwanted reflected sound can distort stereo imagery and degrade the sense of total sound space.

Until recently, most localization research was performed over headphones or in anechoic chambers and the results of that work have not been applicable to normal listening situations with speakers. Our research required that we be able to alter the reverberant characteristics of our room in very selective ways. With the aid of the Crown TEF (Time, Energy, Frequency) analyzer, we have been able to make changes in the placement of sound-absorbent panels on the walls and immediately evaluate the effect of the changes. The TEF analyzer provides us with a higher degree of accuracy and resolution in the detail of room acoustics than has been possible before.

Using the TEF, we have been able to build a room that is selectively anechoic. It is anechoic between the loudspeaker and listener positions, while it appears otherwise reverberant to the listener. This enables us to conduct localization experiments in an environment that can be demonstrated to be free from early reflections, while allowing the subject to experience it as a normal room, not an anechoic chamber. While this environment shares some design goals in common with “Live End—Dead End” studio monitoring rooms, it is intended to be a flexible environment that can be easily altered to fit many different uses.

Fig. 13-16 shows a cutaway view of a small room with a loudspeaker in the corner. Figs. 13-17 through 13-22 show Sonex being added and the effect of each piece of absorption measured. Note that all reflections were removed (in his time window) without total absorption in front of the room. Of course, a control room is much more complex but it shows the direction they are taking.

13.11 Diffusion

Noted authorities have dealt with this problem of diffusion. T. F. W. Embleton stated:

> *In a large, irregular enclosure it is possible, in principle, to have a diffuse sound field, one that consists of a superposition of sound waves traveling in all directions with equal probability. This characteristic ensures that the average energy density (ensemble-average energy density) is the same at all points. If this were actually so, there would be no net flow of power in any direction. Hence, a diffuse sound field never actually exists because there is always a net*

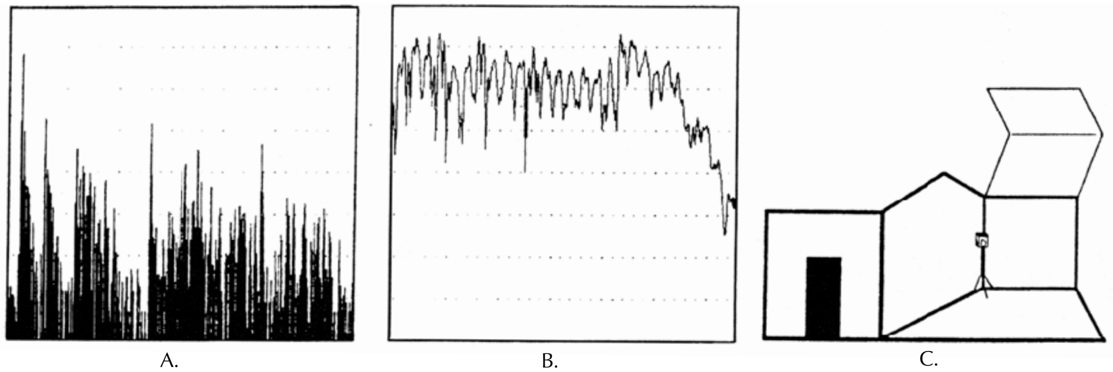

A. B. C.

Figure 13-16. Cutaway of a small room with a loudspeaker in the corner. (*Courtesy Doug Jones.*)

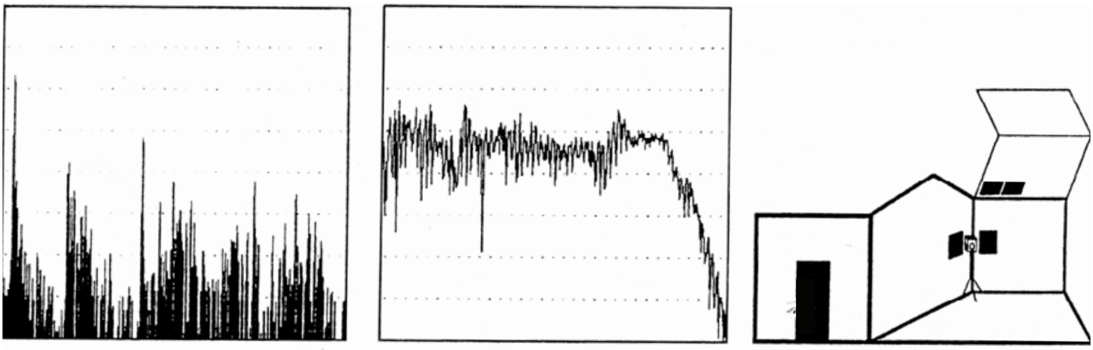

A. B. C.

Figure 13-17. Effect of Sonex being added to the small room in Fig. 13-16. (*Courtesy Doug Jones.*)

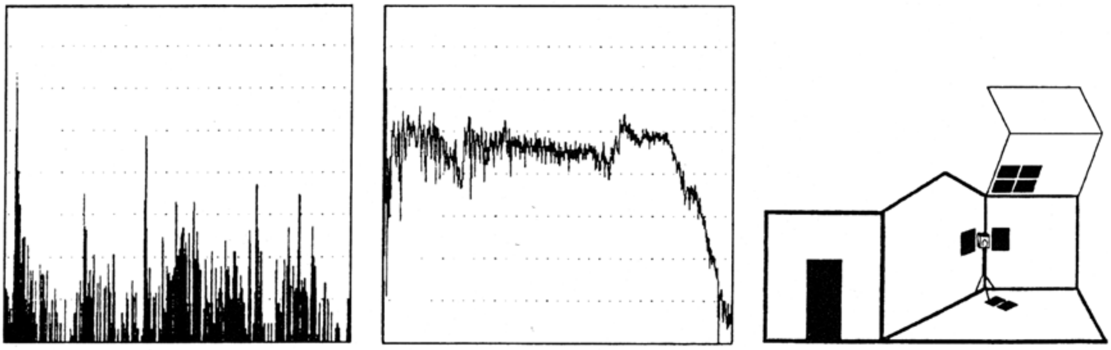

A. B. C.

Figure 13-18. Effect of Sonex being added to the small room in Fig. 13-17. (*Courtesy Doug Jones.*)

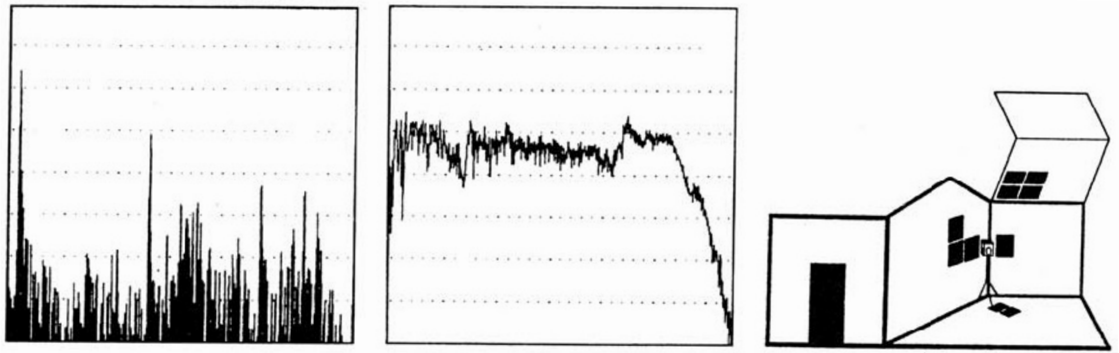

A. B. C.

Figure 13-19. Effect of Sonex being added to the small room in Fig. 13-18. (*Courtesy Doug Jones.*)

> *flow of power away from the source to the places where the energy is ultimately absorbed. Nevertheless, the concept of a diffuse sound field is useful in rooms that are not highly absorptive and where, in addition, the measurement position is neither in the vicinity of the source nor near any small area that is highly absorptive.*

James Moir noted:

> *In an acoustically large room some approach to complete diffusion can exist during the decay period because of the close spacing of the resonant mode frequencies even at the lower audio frequencies.*

Morse and Ingard also noted:

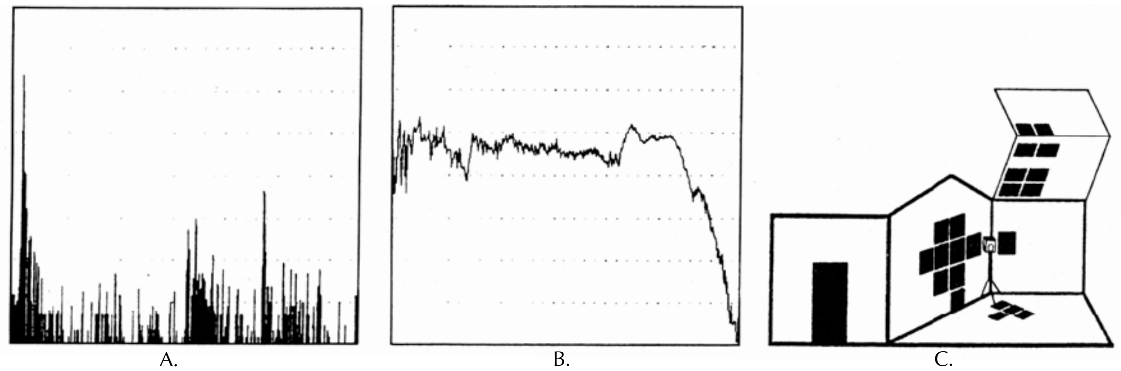


Figure 13-20. Effect of Sonex being added to the small room in Fig. 13-19. (*Courtesy Doug Jones.*)

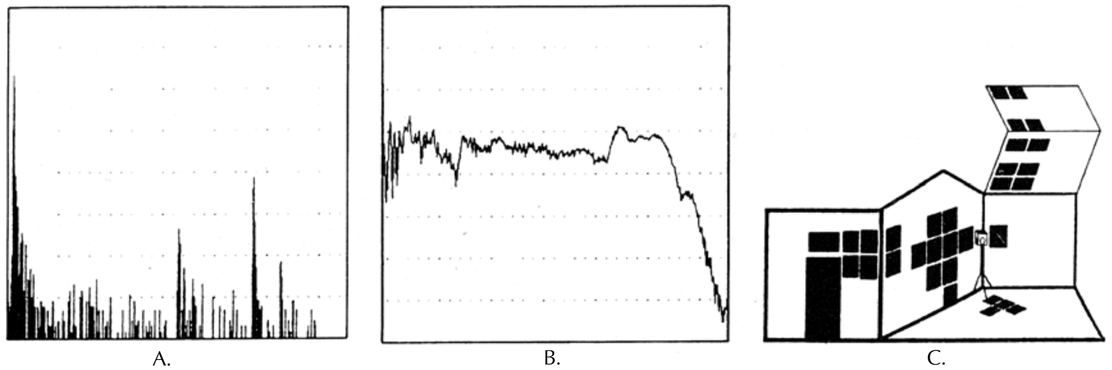


Figure 13-21. Effect of Sonex being added to the small room in Fig. 13-20. (*Courtesy Doug Jones.*)

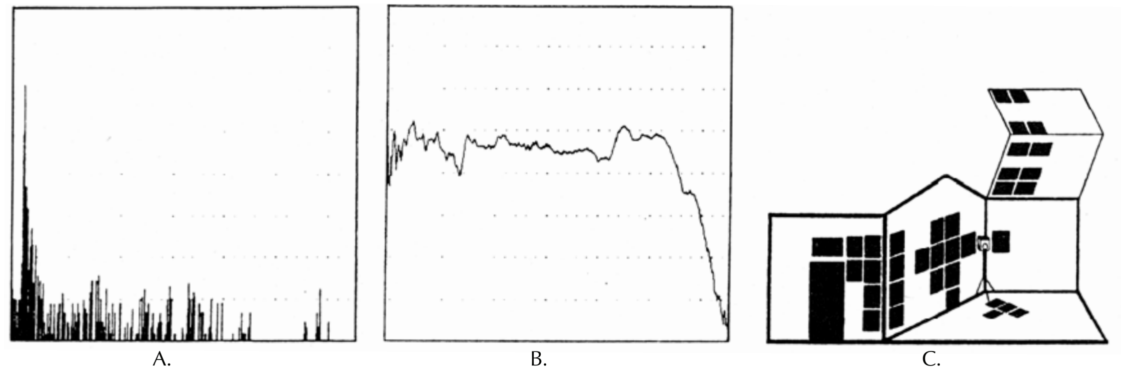
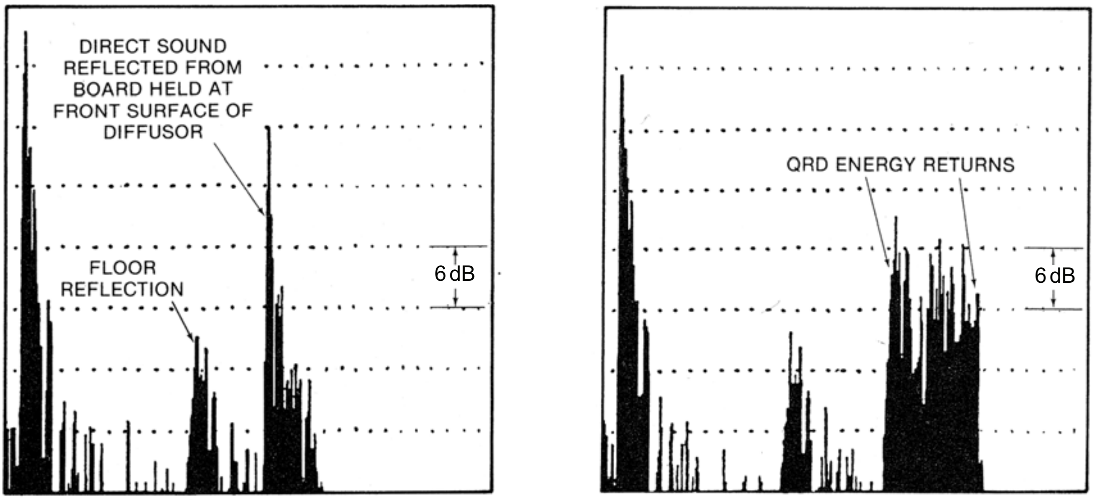


Figure 13-22. Effect of Sonex being added to the small room in Fig. 13-21. (*Courtesy Doug Jones.*)

> *A sound wave is scattered, not only by a solid object, but also by a region in which the acoustic properties of the medium…. Turbulent air scatters, as well as generates sound…. A rough patch on a plane surface scatters, as well as reflects, sound.*

Pick up any book with pretensions to knowledge about recording studios and almost without exception the material on the internal acoustics exhibits an enormous void of accurate or useful information. Implied is that all you have to do is add absorption, with the aid of some devil's apprentice with information from the dark domain, and all is well.

TEF has clearly shown us that the reflection free zone is easy. The tough part is obtaining the optimum diffusion from the "live" end. In fact, the difference in quality of control rooms is the difference in diffusion present at the mixer's ears. The more diffuse and mixed the total sound field at the mixer's ears the better the quality of the sound.



A. B.

Figure 13-23. These TEF measurements were made at the Dallas Sound Lab LEDE Workshop. What the class heard and the measurement they saw assured the future success of diffusors.

13.11.1 Quadratic Residue Diffusors

The first AES paper on LEDE design (D. Davis and C. Davis, 1978) stated that Schroeder's quadratic residue diffusors would be ideal for the diffuse rear wall. Dr. Peter D'Antonio made the Schroeder equations practical. Dr. D'Antonio's hobby is recording. He built an LEDE control room in his basement and used quadratic residue diffusors.

He gave a paper at AES in the fall of 1983 in New York (a poster session). Robert Todrank, who was building an LEDE control room for Jimmy Tarbutton at Acorn Studios in Nashville and was struggling with the diffuse rear wall, attended the session. Todrank built and installed the first diffusors in a commercial control room. Russ Berger, key in the use of the diffusors, was also present, as we were. Berger was to shortly host a Syn-Aud-Con sponsored LEDE workshop at the Dallas Sound Labs. He invited D'Antonio and his diffusors to attend the workshop. The diffusor measurements are shown in Fig. 13-23. The diffusors took the design of the diffuse rear wall out of "art" and made it a predictable success. Subsequently Peter D'Antonio became a manufacturer of diffusors and Russ Berger became the foremost studio control room designer in the world.

13.12 Conclusion

The acoustical properties of rooms are basic to the successful design of sound systems. This chapter on small room acoustics has looked at the ambient noise level L_N in a room, the relationship of direct sound level L_D, to the reverberant sound level L_R (i.e., $L_D - L_R$), and the reverberation time RT_{60}. We'll now turn our attention to how our sound apparatus can be chosen and adjusted to optimize acoustic gain.

Bibliography

P. D'Antonio and J. H. Konnert. "The Reflection Phase Grating Diffusor: Design Theory and Application," *J. Audio Eng. Soc.*, Vol. 32, No. 4 (Apr. 1984), pp. 228-238.

P. D'Antonio, J. H. Konnert, and F. Becker. "The RPG Reflection Phase Grating Diffusor: Experimental Measurements." Paper presented at 76th AES Convention, New York (Oct. 1984). Preprint No. 2158.

P. D'Antonio, J. Konnert, and R. E. Berger. "Control Room Design Utilizing a Reflection Free Zone and Reflection Phase Grating Diffusors: A Case Study." Paper presented at 78th AES Convention, Anaheim, Calif. (May 1985).

C. Davis and D. Davis. "(LEDE)" Live End—Dead End Control Room Acoustics... (TDS) Time Delay Spectrometry...(PZM) Pressure Zone Microphones," *Record. Eng. Prod*. (Feb. 1979).

D. Davis. "Putting It All Together in a Control Room," *Syn-Aud-Con Tech Topics*, Vol. 5 (Apr. 1978).

D. Davis. "Nashville LEDE™ Workshop," *Syn-Aud-Con Tech Topics*, Vol. 12, No. 1 (Fall 1984).

D. Davis and C. Davis. "LEDE and the Diffused Rear Wall," *Syn-Aud-Con Tech Topics*, Vol. 11, No. 7 (1984).

_______. "The LEDE Concept for the Control of Acoustic and Psychoacoustic Parameters in Recording Control Rooms," *J. Audio Eng. Soc*., Vol. 28, No. 9 (Sept. 1980), pp. 585-595.

H. Häas. "The Influence of a Single Echo on the Audibility of Speech," *J. Audio Eng. Soc*., Vol. 20 (Mar. 1972), pp. 145-159.

J. Henry. "On Acoustics Applied To Public Buildings." Paper presented at the American Association for the Advancement of Science, August 1856; later published in the Smithsonian Institution publication.

D. R. Jones. "Designing a Stereo Room for Spatial Hearing Research on Stereo Imagery Using the TEF" (Unpublished).

D. R. Jones, W. L. Martens, and G. S. Kendall. "Optimizing Control Rooms for Stereo Imagery." Paper presented at the Acoustical Society of America, Nashville, Tenn. (Nov. 5, 1985).

H. Kuttruff. *Room Acoustics*. New York: Halstead Press, 1973.

J. Moir. *High Quality Sound Reproduction*. New York: Macmillan, 1958.

C. A. P. Rodgers. "Pinna Transformations and Sound Reproduction," *J. Audio Eng. Soc*., Vol. 29 (1981), pp. 226-234.

M. R. Schroeder. "Progress in Architectural Acoustics and Artificial Reverberation: Concert Hall Acoustics and Number Theory," *J. Audio Eng. Soc*., Vol. 32, No. 4 (Apr. 1984), pp. 194-203.

W. B. Snow. "Application of Acoustical Engineering Principles to Home Music Rooms," *IRE Trans. on Audio* (Nov.-Dec. 1957), pp. 153-159.

R. Todrank. "Incorporating Reflection Phase Grating Diffusors on the Rear Wall to Enhance Spatial Imaging and Stereo Ambiance," *Record. Eng. Prod*., Vol. 15, No. 6 (Dec. 1985), pp. 122-125.

Chapter 14

Designing for Acoustic Gain

by Don Davis

14.1 Maximum Physical Distance 253
14.2 Establishing an Acceptable Signal-to-Noise Ratio (*SNR*) 253
14.3 Establishing an *EAD* 254
14.4 Needed Acoustic Gain (*NAG*) 254
14.5 The Number of Open Microphones 255
14.6 The Feedback Stability Margin 255
14.7 Calculating Potential Acoustic Gain 256
    Acoustic Gain Parameters 258
    Effect of Directional Devices on Acoustic Gain 258
14.8 Obtaining ΔD_x Values 258
14.9 Measuring Acoustic Gain 259
14.10 Achieving Potential Acoustic Gain 259
14.11 Limiting Parameters in Sound Reinforcement System Design 260
14.12 How Much Electrical Power Is Required? 260
14.13 Finding the Required Electrical Power (*REP*) 261
    Loudspeaker Efficiency 262
    Conversions 262
    100% Efficiency Values 263
14.14 Summary 263
Bibliography 263

Years of experience with acoustic gain equations have demonstrated their usefulness and accuracy in identifying the maximum acoustic level possible in a sound reinforcement system. Direct sound levels at the furthest listener have been found to be of greater importance than the total sound level.

A number of factors must be present at the listener's position for easily understood communication to take place:

1. The sound must be sufficiently loud, and it must be at least 25 dB above the ambient noise level at midfrequencies (2000 Hz) in rooms with $RT_{60} \geq 1.6\,\text{s}$).
2. The sound must reasonably approximate the same spectrum shape as that produced by the talker or other source.
3. The sound must reasonably approximate a ratio of direct-to-reverberant sound within the rules for acceptable articulation loss. Temporal misalignment of arrays with the resultant changes in polar response (lobing) is the most common cause of an unexpectedly poor direct-to-reverberant ratio (D/R).

14.1 Maximum Physical Distance

Using a sample acoustic environment where:
$V = 500{,}000\,\text{ft}^3$,
$S = 42{,}500\,\text{ft}^2$,
$\bar{a} = 0.206$,
$RT_{60} = 2.5\,\text{s}$,
and testing it for the maximum physical distance a talker and a listener could stand apart and easily be heard and understood without a sound system, we would find that we have indeed fulfilled the three criteria mentioned above. This physical distance is typically 6–10 ft in the environment for this sample. For example, if a weak talker produced 65 dBA at 2 ft in a very quiet room and the ambient noise was 28 dBA, then:

$$L_N + SNR = L_{req} \tag{14-1}$$

where,
L_N is the noise level,
SNR is the signal-to-noise ratio,
L_{req} is the required level.

$$28 + 25 = 53\text{ dB}$$

$$65\text{ dBA talker} - 53\text{ dBA required} = 12\text{ dB}$$

$$2\text{ ft} \times 10^{\frac{12\text{ dB}}{20\text{ dB}}} = 8\text{ ft}$$

This example shows that a distance of 8 ft from the talker is the maximum physical distance at which a listener can stand and be sure of hearing clearly if no sound system is present. If we use a normal male voice level of 71 dB at 2 ft, then the maximum physical distance would be:

$$71\text{ dBA} - 53\text{ dBA} = 18\text{ dBA}$$

$$2\text{ ft} \times 10^{\frac{18\text{ dB}}{20\text{ dB}}} = 16\text{ ft}$$

Fig. 14-1 allows these calculations to be made at a glance. For example, our normal voice of 71 dBA at 2 ft becomes 83 dBA at 0.5 ft. Going up the normal voice line to 53 dB results in an $EAD \approx 16$ ft. "Weak voice" and "normal voice" are relatively self-explanatory. The expected voice level is a very real effect that can be relied upon in marginal communication circumstances or that must be carefully avoided in the case of speech privacy systems, for example.

If a talker and a listener were 10 ft apart and we raised the noise level by means of a speech privacy system to an L_P of approximately 58 dB overall, then the talker would involuntarily raise his or her voice to overcome the noise. If we kept the noise at 55 dBA, then masking would be effective at distances of 12–15 ft from the talker, who would continue to converse at normal voice level. A talker can raise Q slightly by using cupped hands in front of the mouth, megaphone fashion; a listener can raise Q by using cupped hands behind the ears.

The limit shown for amplified speech is not due to an inability to amplify but because these levels represent dangerous L_P conditions for the listener's ears. Fig. 14-1 outlines the parameters of the possible and impossible sound system solutions and is simple enough for laymen to understand during presentations.

Having arrived at the maximum physical distance between a talker and a listener with no sound system present, we now adopt that distance as our goal for a successful sound system. We want the equivalent acoustic distance (EAD) to be established at the farthest listener (D_2), even though this listener may be over 100 ft away.

14.2 Establishing an Acceptable Signal-to-Noise Ratio (*SNR*)

Establishing an acceptable signal-to-noise ratio, *SNR*, is a major factor in achieving intelligibility. We typically want at least 25 dB of *SNR* in the 2000 Hz octave band. Establishing such a ratio can

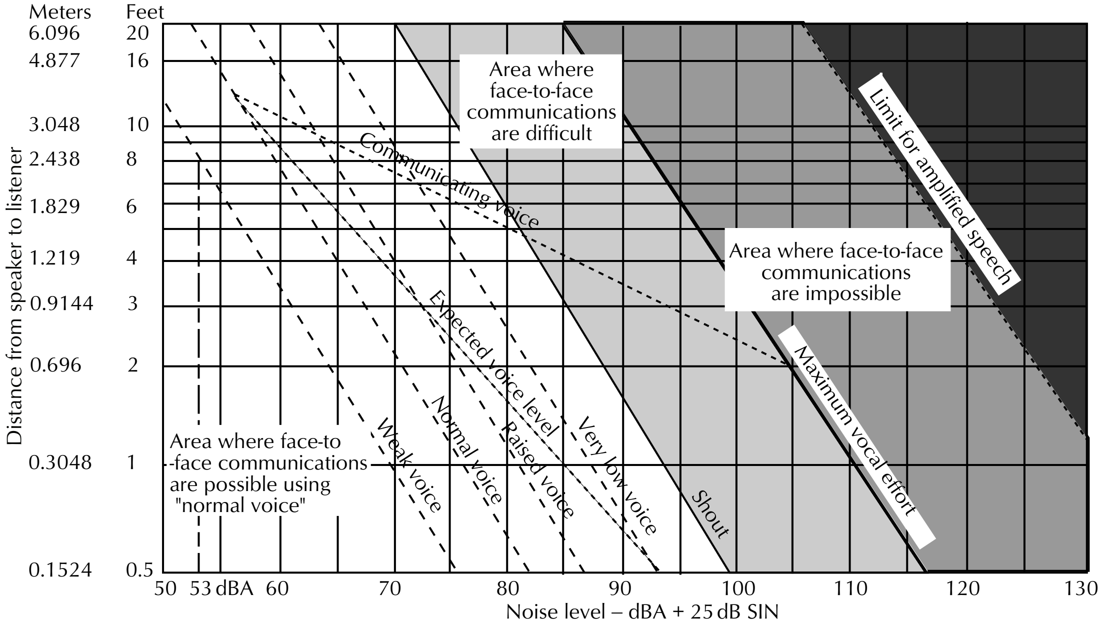


Figure 14-1. Nomograph for finding the *EAD*. (*Courtesy John Webster.*)

be a problem itself. Consider, for example, one segment of a large audience area that is affected by a noisy air handler so its *EAD* is different than the rest of the room. Where the level of the ambient noise field is relatively constant throughout the listening area, we can use the equivalent acoustic distance (*EAD*) everywhere.

14.3 Establishing an *EAD*

The *EAD* requires that we meet the following criteria:

1. Establish at the listener's ears (at 125 ft, for example) the same L_P, via the sound system, that would have been heard at the maximum physical distance (in the sample case, the same L_P that would have been heard 8 ft from the talker).
2. Establish at the listener's ears, via the sound system, the same spectrum shape that would have been heard at the maximum physical distance.
3. Establish at the listener's ears, via the sound system, a ratio of direct-to-reflected sound that does not deteriorate the articulation loss of consonants ($\%AL_{CONS}$) by more than 15%. In other words, the equivalent acoustic distance establishes a set of conditions that are the same as the maximum physical distance without a sound system at some much greater distance from the source through the use of a sound system.

14.4 Needed Acoustic Gain (*NAG*)

Using the standard sound system notation, we can look at Fig. 14-2 and see that we have a real listener at some distance from the talker (D_0). The talker is a given distance from the microphone (D_s). To find the needed acoustic gain (*NAG*) find the attenuation over distance D_0 minus the attenuation over distance *EAD*. For example, we are outdoors and the inverse-square-law level change applies. The following distances are known:

$D_S = 2\,\text{ft}$,

$EAD = 8\,\text{ft}$,

$D_0 = 128\,\text{ft}$.

We can then write:

$$NAG = 20\log\frac{D_0}{D_s} - 20\log\frac{EAD}{D_s} \qquad (14\text{-}2)$$

or simply:

$$NAG = 20\log D_0 - 20\log EAD \qquad (14\text{-}3)$$

which in the example is:

$$\begin{aligned} NAG &= 20\log 128 - 20\log 8 \\ &= 24\ \text{dB} \end{aligned}$$

where,

EAD is the equivalent acoustic distance,

D_0 is the distance from talker to farthest listener,

D_s is the distance from talker to microphone.

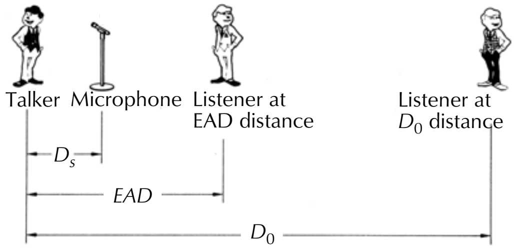


Figure 14-2. Distances involved in determining *NAG*.

We can also calculate *NAG* the long way. If the talker produced 70 dBA at a microphone 2 ft away, then the talker would produce 64 dBA at 4 ft and 58 dBA at 8 ft (6 dB loss with doubling of distance). Continuing on, at 16 ft the level would be down to 52 dBA, at 32 ft it would be 46 dBA, at 64 ft it would be 40 dBA, and at 128 ft it would be 34 dBA. The L_P we need is that at 8 ft, or 58 dBA. Then:

$$\begin{aligned} NAG &= \text{dB SPL at } 8\text{ ft} - \text{dB SPL at } 128\text{ ft} \\ &= 58\text{ dB} - 34\text{ dB} \\ &= 24\text{ dB} \end{aligned}$$

We thus need 24 dB of acoustic gain to have an *EAD* of 8 ft at 128 ft.

14.5 The Number of Open Microphones

If we raise two microphones to the same level in a reinforcement system, we will have to reduce the overall gain 3 dB to avoid the system going into feedback (remember how decibels combine). Since each microphone is sampling the sound field, every time you double the number of open microphones (*NOM*), we will have to lower the gain 3 dB, Fig. 14-3. Therefore the loss in gain caused by more than one open microphone is:

$$NOM\text{ (in dB)} = 10\log NOM. \qquad (14\text{-}4)$$

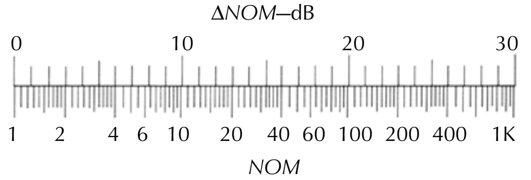


Figure 14-3. Chart relating *NOM* and *NOM* in decibels.

If we decide to operate a sound system with more than one open microphone, we must add the decibels we will lose through having these extra microphones active to the gain required so *NAG* now becomes:

$$NAG = \Delta D_0 - \Delta EAD + 10\log NOM \qquad (14\text{-}5)$$

14.6 The Feedback Stability Margin

In a paper that turned out to be remarkably ahead of its time, William B. Snow, in the April 1955 *AES Journal*, described the detrimental effects of operating a sound system too near the acoustic feedback point. The paper, *Frequency Characteristics of a Sound Reinforcing System*, contained a set of superlative illustrations made by Snow on his level recorder. These illustrations are used here to point out the main features of this pioneer paper.

Fig. 14-4 shows the effect on the response of the system as it approaches unity gain. The 10, 20, 30, and 40 markings are arbitrary settings on an amplifier's gain control. For example, the curve marked 20 was 15 dB below feedback. Note that, even at 25 dB below feedback (30 on the chart), significant irregularities are evident. Note also that large corrections are not necessary and that "bumps" corrected by less than 1 dB at –25 dB below feedback will not grow into bumps when the gain is again raised. When a sound system is unequalized, at least 12 dB of feedback stability margin (*FSM*) is required. When carefully equalized, 6 dB of *FSM* is adequate to ensure a stable system free of spurious regenerative sounds.

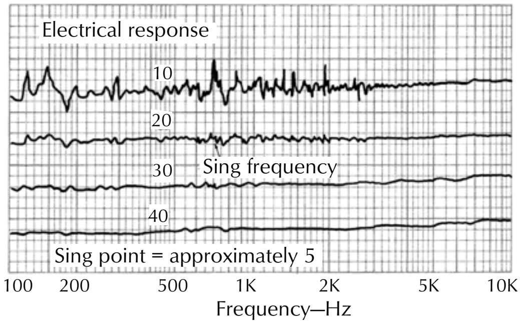


Figure 14-4. Response of a sound system as it operates with gain. Sing frequency is 5 dB. Therefore, 30 is 25 dB below feedback.

Boost devices require excessive *FSM* to be safe to use. Fig. 14-5 illustrates the nonlinear behavior of such minor aberrations in response when bass boost is used too near regeneration. Fig. 14-6 illustrates the same effect at the opposite end of the spectrum when treble boost is used.

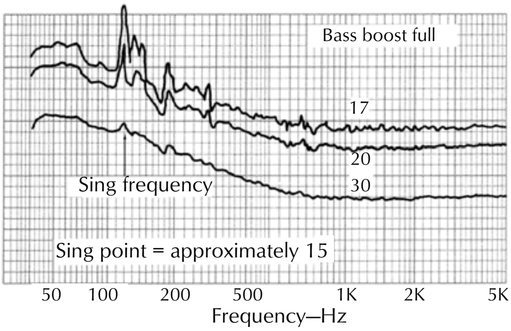


Figure 14-5. Nonlinear response of system using bass boost.

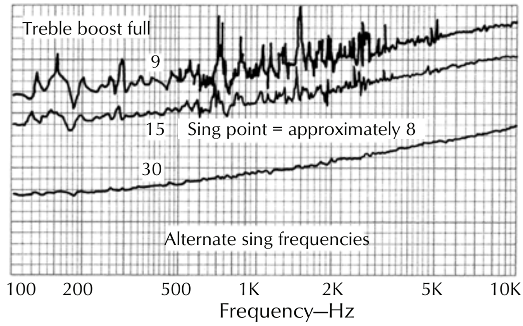


Figure 14-6. Nonlinear response of system using treble boost.

Fig. 14-7 explains a still different phenomenon. As a sound system is brought near feedback, not only does the amplitude response behave nonlinearly but the transient response also does so. The natural reverberation time of the room is magnified by the sound system. In Snow's demonstration room, RT_{60} = 0.65 s. When the sound system was just below feedback, RT_{60} was 2.7 s, a multiplication factor of 4.2 times:

$$RT_{60} = \frac{60}{x} \tag{14-6}$$

where,

x is the decay rate in dB/s.

Fig. 14-8 shows an example of the amplitude non-linearity as seen on a ⅓ octave analyzer.

Many early experiments in the art of equalization mistook the regenerative amplitudes for the true amplitude and used filters that were set too deep. Although any small irregularity in response can be magnified as shown here, some won't be for a given position of microphone and loudspeaker because of the total room system phase response. Consequently, it becomes apparent that the free-field response of desirable transducers for sound-reinforcement work must exhibit relatively smooth changes in response without peaks or dips that have rapid slope rate changes.

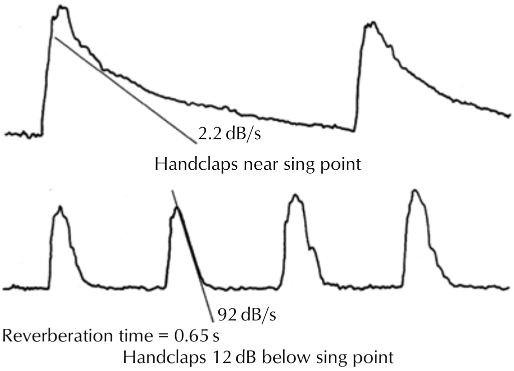


Figure 14-7. Nonlinear behavior of amplitude and transient response.

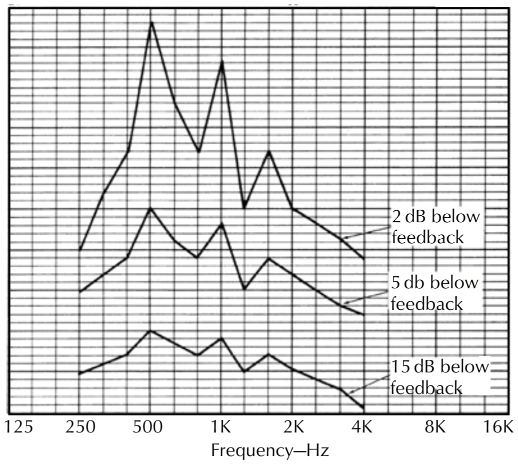


Figure 14-8. Regenerative swelling of normal response 15 dB, 5 dB, and 2 dB below feedback.

Snow (1955), Davis (1967), Mankovsky (1971), and Yamamoto (1971) have found a *FSM* of 6 dB necessary. We thus need to add to our *NAG* this 6 dB *FSM*. We now can write a general formula for finding the needed acoustic gain of a sound-reinforcement system:

$$NAG = \Delta D_0 - \Delta EAD + 10\log NOM + 6\text{ dB FSM} \tag{14-7}$$

14.7 Calculating Potential Acoustic Gain

Assuming that inverse-square-law level change will serve in an outdoor situation and indoors for L_D, we can construct the set of sound system parameters shown in Fig. 14-9. In conjunction with these calculations, we can use the nomograph in Fig. 14-10.

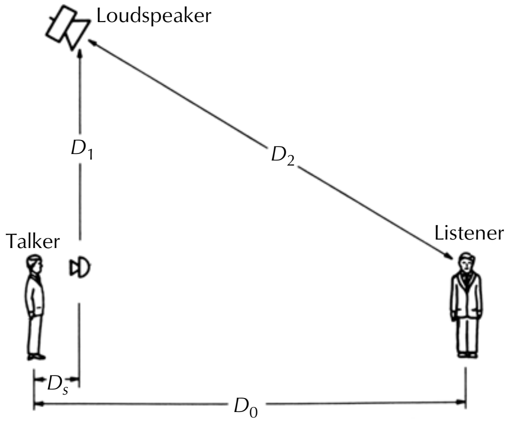


Figure 14-9. Basic parameters of a single-source system.

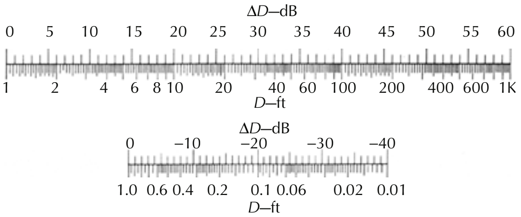


Figure 14-10. Calculation of relative changes in level with distance (inverse square law).

If the required gain ($\Delta D_0 - \Delta EAD$) is 24 dB + 6 dB FSM and one microphone is open, substituting in the equation gives:

$$\begin{aligned} NAG &= 24\text{ dB} + 0\text{ dB} + 6\text{ dB} \\ &= 30\text{ dB} \end{aligned}$$

If we assign the following values to the parameters in Fig. 14-9, we can find their dB equivalents by using Fig. 14-10:

$D_0 = 128$ ft,
$\Delta D_0 = 42$ dB,
$D_S = 2$ ft,
$\Delta D_s = 6$ dB,
$D_1 = 45$ ft,
$\Delta D_1 = 33$ dB,
$D_2 = 90$ ft,
$\Delta D_2 = 39$ dB.

We can write the following equation for potential acoustic gain (*PAG*):

$$PAG = \Delta D_0 + \Delta D_1 - \Delta D_s - \Delta D_2 \qquad (14\text{-}8)$$

where,
D_0 is the distance from talker to farthest listener,
D_1 is the distance from the microphone to the loudspeaker,
D_s is the distance from talker to microphone,
D_2 is the distance from the loudspeaker to the farthest listener.

In our example:

$$\begin{aligned} PAG &= 42\text{ dB} + 33\text{ dB} - 6\text{ dB} - 39\text{ dB} \\ &= 30\text{ dB} \end{aligned}$$

Since *PAG* = *NAG*, we have sufficient acoustic gain. We can write this another way:

$$PAG - NAG = 0 \qquad (14\text{-}9)$$

If we actually write the equations in this manner, we discover:

$$\begin{aligned} &\Delta D_0 + \Delta D_1 - \Delta D_s - \Delta D_2 = \\ &\Delta D_0 - \Delta EAD + 10\log NOM + 6 \end{aligned} \qquad (14\text{-}10)$$

or

$$\begin{aligned} &\Delta D_0 - \Delta D_0 + \Delta D_1 + EAD - \Delta D_s - \Delta D_2 \\ &- 10\log NOM - 6 = 0 \end{aligned} \qquad (14\text{-}11)$$

The ΔD_0s cancel, and we arrive at a most useful general formula:

$$\Delta D_1 + \Delta EAD - \Delta D_s - \Delta D_2 - 10\log NOM - 6 = 0 \qquad (14\text{-}12)$$

Our original requirement was a *NAG* of 30 dB with an *EAD* of 8 ft and a D_s of 2 ft. We could have written the general equation in the following manner to find ΔD_2:

$$\Delta D_1 + \Delta EAD - \Delta D_s - 10\log NOM - 6 = \Delta D_2 \qquad (14\text{-}13)$$

Because signal delay would become a factor if we made D_1 greater than 45 ft, we chose 45 ft as a limit on D_1:

$$\text{Optimum } D_1 \geq D_C < 45\text{ ft} \qquad (14\text{-}14)$$

We could now write:

$$\Delta 45\text{ ft} + \Delta 8\text{ ft} - \Delta 2\text{ ft} - 10\log 1 - 6 = \Delta D_2$$

or

$$33\text{ dB} + 18\text{ dB} - 6\text{ dB} - 0\text{ dB} - 6\text{ dB} = 39\text{ dB}$$

Looking on the dB part of the scale in Fig. 14-10, we find that 39 dB is equivalent to 90 ft.

Now that we have seen how our values are found, we can simplify the process by removing the Δ operator and using the following ratio equations:

$$\text{Max } D_s = \frac{D_1 \times EAD}{2D_2\sqrt{NOM}} \tag{14-15}$$

$$\text{Min } D_1 = \frac{2D_s \times D_2 \times \sqrt{NOM}}{EAD} \tag{14-16}$$

$$\text{Max } D_2 = \frac{D_1 \times EAD}{2D_s\sqrt{NOM}} \tag{14-17}$$

$$\text{Min } EAD = \frac{2D_s \times D_2 \times \sqrt{NOM}}{D_1} \tag{14-18}$$

$$\text{Max } NOM = \frac{(D_1 \times EAD)^2}{(D_s \times 2D_2)^2} \tag{14-19}$$

All parameters are now in feet or meters, except *NOM*, which is the number of open microphones.

14.7.1 Acoustic Gain Parameters

Using Figs. 14-9 and 14-10, we have already determined that the maximum D_2 possible is 90 ft. Why is it the maximum distance? Let's examine what happens as we change each of these basic parameters, one at a time.

If D_1 is increased, what happens to the acoustic gain? Since the loudspeaker and microphone are separated further, the gain can be turned up more before the sound from the loudspeaker reaches the microphone at the same level as the talker's voice (unity gain). In fact, let's look at the series of level changes that makes a sound system necessary and allows it to work at all.

If, as in our example for the calculation of *NAG*, we again assume that the talker generates 70 dBA at a microphone 2 ft away, we can also assume that the loudspeaker can deliver 70 dBA at the microphone just as feedback begins. If we work backward with inverse-square-law level change, we find that at 4 ft, the loudspeaker is providing:

$$70\text{ dBA} + 20\log\frac{45\text{ ft}}{4\text{ ft}} = 91\text{ dBA}$$

Therefore, increasing D_1 increases the acoustic gain until D_c acts as a limit.

Going away from the loudspeaker in the direction of the listener (D_2), we see that at 8 ft we have 85 dBA, at 16 ft we have 79 dBA, at 32 ft we have 73 dBA, at 64 ft we have 67 dBA, and at 90 ft we have 64 dBA. Note that 64 dBA is just 6 dB greater than 58 dBA. Now, if D_2 is increased, the level beyond 90 ft would also decrease; therefore, increasing D_2 lowers the apparent acoustic gain.

If D_s is increased, the result is obvious. Any time we move farther away from the microphone, we lose apparent acoustic gain.

Finally, what happens if D_0 is increased? The level change between the talker and the listener increases. Since the sound arriving from the loudspeaker has remained the same (remember we are changing only one parameter at a time), the apparent gain has increased.

Eq. 14-7, $PAG = \Delta D_0 + \Delta D_1 - \Delta D_s - \Delta D_2$, shows by means of the signs (+ or −) which parameters will increase apparent gain if they are increased (all plus signs do this) and which parameters will decrease apparent gain if they are increased (all minus signs do this).

Note that the absolute acoustic gain of the system is determined by the true acoustic separation between the microphone and the loudspeaker (D_1). All other parameters change the apparent acoustic gain (as observed by the listener).

14.7.2 Effect of Directional Devices on Acoustic Gain

To find the effect of directional devices on acoustic gain, do the following:

1. Find *PAG* for an omnidirectional source (PAG_{omni}).
2. Take the difference in dB between the level on the microphone polar plot at the angle toward the talker and the angle toward the loudspeaker, $MS_\angle$.
3. Take the difference in dB between the level on the loudspeaker's polar plot at the angle toward the listener and the angle toward the microphone, $SM_\angle$.
4. $M_e = 10^{\left(\frac{MS_\angle + SM_\angle}{20}\right)}$.
5. $PAG_{omni} + (MS_\angle + SM_\angle)$ = Total gain (free field), Fig. 14-11.

14.8 Obtaining ΔD_x Values

In obtaining ΔD_x values (distances converted into relative levels) for use in acoustic gain equations, the following techniques are the basic ones:

1. Use inverse-square-law level change for free-field conditions:

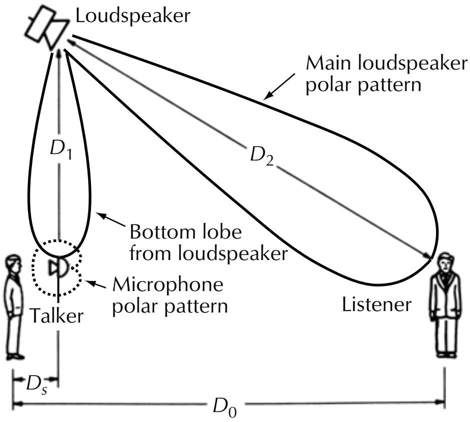


Figure 14-11. Sound radiation patterns.

$$\Delta D_x = 20\log D_x \tag{14-20}$$

This automatically makes the reference distance unity and since it is nondimensional, either U.S. or SI lengths may be used.

2. Use the Hopkins-Stryker equation as a relative level generator for reverberant spaces, especially when specific dB per doubling of distance beyond D_c are known.

14.9 Measuring Acoustic Gain

It is one thing to calculate gains at the drawing-board stage of a construction project. The real thrill comes when, after having written your calculation into a specification, many months later you actually measure the acoustic gain of the sound system after installation. The architects, engineers, owners, and other interested parties know your prediction. When your actual measurements come within ±2 dB of your calculation, which has happened in literally hundreds of jobs, these professionals have no choice but to regard you as a fellow professional.

Measuring the acoustic gain of a finished sound system is described below and in Fig. 14-12:

1. With the sound system turned off and the room made as quiet as possible (air conditioning or heating turned off), adjust the test amplifier (use pink noise input) to give a 75–80 dBA reading over the test loudspeaker at the sound system microphone (substitute the sound level meter temporarily for the microphone). The microphone should be placed at its design position.
2. Carry the sound level meter to the farthest D_2 position and measure the level from the test loudspeaker. Exercise care that the signal arriving from the test speaker is at least 6 dB greater than the ambient noise reading. Record this reading in dBA.
3. Turn on the sound system (use the same test loudspeaker feeding a 75–80 dBA signal into the microphone) and adjust below self-sustaining feedback (the system will ring, but upon cutting off the signal, the feedback should not continue).
4. Again read the sound level meter at the farthest D_2 position and record in dBA.
5. The reading taken with the sound system on minus the reading taken with the sound system off equals the total acoustic gain. The total acoustic gain should be within ±2 dB of the calculated *PAG*.

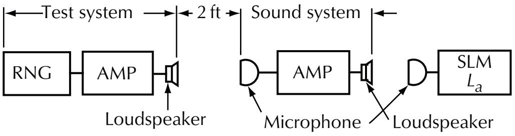


Figure 14-12. Method of measuring acoustic gain.

14.10 Achieving Potential Acoustic Gain

The potential gain figures illustrated in this chapter depend on the performance of room-sound-system equalization (Chapter 24 *Sound System Equalization*) to ensure unity gain at all frequencies of interest.

Equally important to achieving potential acoustic gain is freedom from misaligned loudspeakers and focused energy coming back into the microphone. Either of these two problems will seriously deteriorate intelligibility and gain.

A piece of sound absorbing material moved about the microphone, the loudspeaker, and a listener in the audience seating area will enable you to determine where the problem originates. Once the problem is isolated, the solution is often easy.

Be sure to measure what sound is emitting from the back and bottom of the loudspeaker as well as from the front. Remember that bass is omnidirectional, Fig. 14-13. Frequency curve A is what the audience would have heard had the microphone not heard Curve B from the back and bottom of the loudspeaker. Enormous equalization was required in the 200 Hz to 500 Hz region, totally destroying male voice intelligibility.

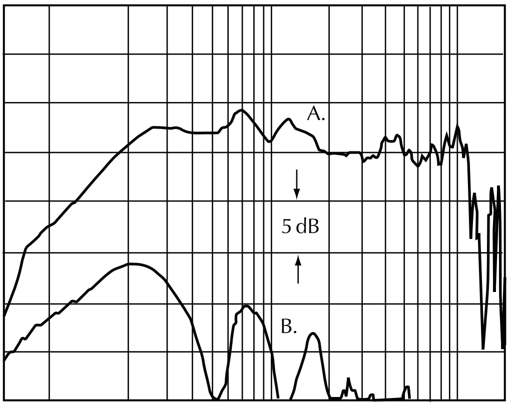


Figure 14-13. Sound measured from front and back of speaker—40–12,000 Hz.

14.11 Limiting Parameters in Sound Reinforcement System Design

Certain choices of parameters limit the choice of approach to the system design. For example, if we wish to use a single-source system (and all other room parameters allow it) because of its inherently more economical approach and ability to blend acoustically with the live talker, then we cannot have an *EAD* of less than twice D_s. For example, if we must have a D_s of 2 ft, then the *EAD* cannot be less than 4 ft:

$$EAD \geq 2D_s \tag{14-21}$$

Naturally, the other limit on *EAD* is that it not exceed D_c:

$$EAD \leq D_c \tag{14-22}$$

This limitation ($EAD \geq 2D_s$) for a single-source sound system is on account of the *FSM*. Fig. 14-14 illustrates the limiting case. When *EAD* must equal D_s, a distributed sound system becomes mandatory. When $D_s = EAD$, then D_1 must ≥ $2D_s$, see Fig. 14-15. Fig. 14-16 shows how D_1 and D_2 are handled in calculating the acoustic gain of a distributed system.

14.12 How Much Electrical Power Is Required?

While the calculation of acoustic gain is pertinent only to sound-reinforcement systems, the same technique can be applied to finding out how to achieve a desired level at some distance point from a sound source. Rock groups, especially, make heavy

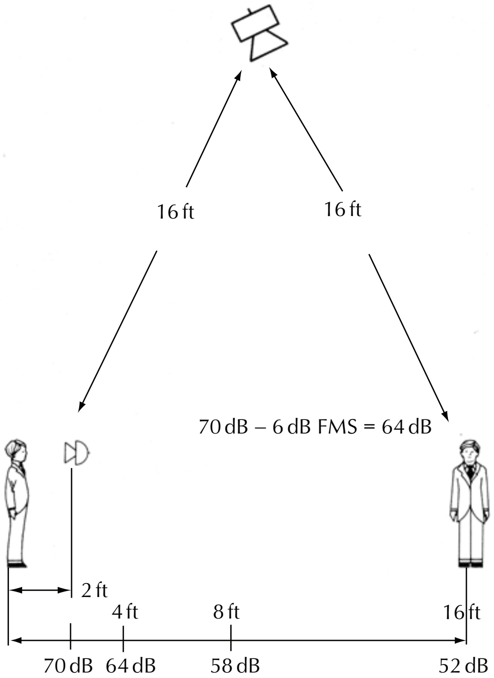

16 ft
16 ft
70 dB – 6 dB FMS = 64 dB
2 ft
4 ft
8 ft
16 ft
70 dB
64 dB
58 dB
52 dB

Figure 14-14. Effect of *FSM* on acoustic gain.

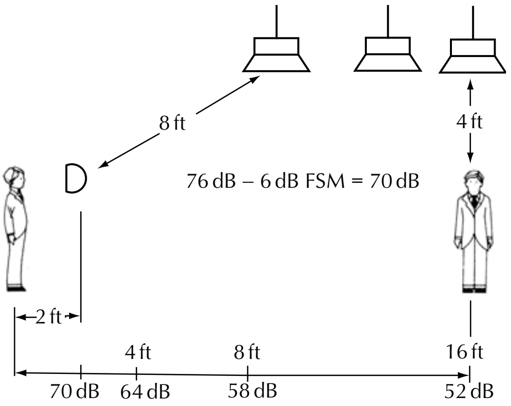

8 ft
4 ft
76 dB – 6 dB FSM = 70 dB
2 ft
4 ft
8 ft
16 ft
70 dB
64 dB
58 dB
52 dB

Figure 14-15. Requirement if D_s is equal to *EAD*.

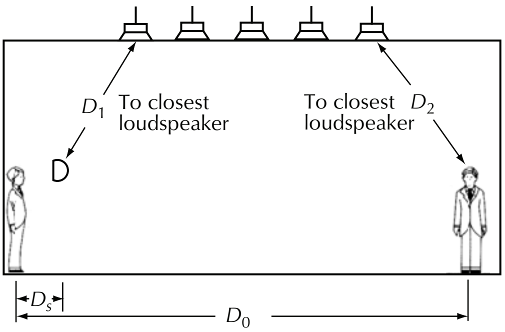

D_1
To closest loudspeaker
To closest D_2 loudspeaker
D_s
D_0

Figure 14-16. Basic parameters of a distributed system.

demands on the power-handling capabilities of sound equipment. The ability to predict accurately what they will get for their investment in different locations can be useful. When you have some definite acoustic sound-pressure-level (L_P) goal in mind at some given distance (D_2) from the loudspeaker, you need to know two important details:

1. The EIA sensitivity rating of the loudspeaker, measured at 30 ft on axis when the loudspeaker is fed an input signal of 0.001 electrical watt, Fig. 14-17.
2. The acoustic level change and attenuation between the loudspeaker as measured at its reference distance for sensitivity and the farthest listener position.

Once the desired acoustic level at the farthest listener has been determined by actual measurement, experience, or calculation, then the desired acoustic level plus the acoustic level change over distance D_2 equals the 30 ft rating the loudspeaker will be required to produce. EIA rating + 30 ft dB rating = power required in dBm.

Sensitivity Ratings Currently in Use

1. *N* dB at 4 ft from 1.0 W input
2. *N* dB at 30 ft from 0.001 W Input
3. *N* dB at 1.0 m from 1.0 W input
4. *N* W needed to produce 1.0 Pa at 1.0 m

European Sensitivity Rating

While in Europe (March 1982) we were told that they are rating loudspeaker sensitivity as: Watts needed for a sound pressure of 1 Pa at 1 m (W/Pa/m)

Example

Given a loudspeaker rated at 99 dB at 4 ft for an electric power Input of 1 W, then:

$$10^{\frac{\left(94 - 99 + 20\log\frac{4}{3.28}\right)}{10}} = 0.21\ \text{W/Pa/m}$$

There are 3.280839895 ft/m. That is, a power of 0.21 W will produce a sound pressure of 1.0 Pa at 1 m. (1 Pa is a sound pressure level of 94 dB.)

Figure 14-17. Sensitivity ratings.

14.13 Finding the Required Electrical Power (*REP*)

Every dB of gain achieved by careful design and equalization requires power to support that gain at the output of the sound system. The advent of acoustical gain formulas and their optimization via equalization marked the onset of more powerful sound systems than had previously been the case.

The electrical power required to produce a pressure level L_P at a distance D_2 is given by:

$$REP = 10^{\frac{[(L_P + 10) + (20\log D_2) - 20\log D_{REF} - L_{SENSI}]}{10}} \times W_{REF} \tag{14-23}$$

For example, to produce an L_P = 90 dB at a distance of 128 ft:

$$REP = 10^{\frac{[(90 + 10) + (20\log 128) - 20\log 30 - 51.5]}{10}} \times 0.001\ \text{W}$$

$$REP = 1228\ \text{W}$$

In our example we have chosen a medium efficiency loudspeaker that has a power rating of 50 W. This leaves us with two choices:

1. Select a more efficient loudspeaker.
2. Accept a lower level at the listener.

If we choose the first alternative, larger size, higher cost, we can find the new power level by:

$$REP = 10^{\frac{[(90 + 10) + (20\log 128) - 20\log 30 - 61.5]}{10}} \times 0.001\ \text{W}$$

$$REP = 51\ \text{W}$$

If we embrace choice two (2) we can find the listener's maximum level $L_{P(MAX)}$ by:

$$L_{P(MAX)} = dBm + L_{SENSI} - 20\log D_2 + 20\log D_{REF} \tag{14-24}$$

$$\begin{aligned} L_{P(MAX)} &= 47^{*}\ \text{dBm} + 51.5 - 20\log 128 + 20\log 30 \\ &= 85.9\ \text{dB} \end{aligned}$$

*50 W

The first loudspeaker was a medium efficiency loudspeaker. The second was a very high efficiency professional sound horn system. Imagine this same requirement using a "home" type low efficiency system with an L_{EIA} of 41.5 dB

$$REP = 10^{\frac{[(90 + 10) + (20\log 128) - 20\log 30 - 41.5]}{10}} \times 0.001\ \text{W}$$

$$REP = 12{,}882\ \text{W}$$

All three of these examples actually exist in the market place and illustrate the necessity to monitor both sensitivity ratings and efficiency values.

14.13.1 Loudspeaker Efficiency

You will often hear the specious argument, “Watts are cheap, so don’t worry about loudspeaker efficiency.” This might be partially true in very small apartment living rooms but it is not true in professional sound work. To find the % efficiency when the L_{EIA} and Q are known, use:

$$\%\,Effici = 10^{\left[\frac{L_{EIA} - 10\log Q - 59.78}{10}\right]} \times 100\% \quad (14\text{-}25)$$

For a Q of 7:

$$\begin{aligned}\%\,Effici &= 10^{\left[\frac{51.5 - 10\log 7 - 59.78}{10}\right]} \times 100\% \\ &= 2.1\%\end{aligned}$$

For a Q of 20:

$$\begin{aligned}\%\,Effici &= 10^{\left[\frac{65.5 - 10\log 20 - 59.78}{10}\right]} \times 100\% \\ &= 18.7\%\end{aligned}$$

For a Q of 3:

$$\begin{aligned}\%\,Effici &= 10^{\left[\frac{41.5 - 10\log 3 - 59.78}{10}\right]} \times 100\% \\ &= 0.5\%\end{aligned}$$

To find the EIA sensitivity from any sensitivity rating use:

$$L_{EIA} = L_{SENSI} + 20\log\left(\frac{D_{MEAS}}{D_{REF}}\right) + 10\log\left(\frac{W_{REF}}{W_{MEAS}}\right) \quad (14\text{-}26)$$

14.13.2 Conversions

- 4 ft = 1.219 m.
- 30 ft = 9.144 m.
- 3.281 ft = 1.0 m.

For an omnidirectional spherical radiator of 0.282 m (surface area = 1.0 m^2), 1.0 W produces an L_W, L_I, and L_P numerically equal to 120 dB.

Ratings

- *%Effici* for 1.0 m, 1.0 W sensitivity.
- *%Effici* for 4.0 ft, 1.0 W sensitivity.
- *%Effici* for 30 ft, 1.0 W sensitivity.
- *%Effici* for W/Pa/m sensitivity.

Calculations

For 1.0 W, 1.0 m:

$$\%\,Effici = \left(\frac{1}{Q} \times 10^{\left(\frac{SENSI}{10} - 10.9\right)}\right) \times 100 \quad (14\text{-}27)$$

For 1.0 W, 4 ft:

$$\%\,Effici = \left(\frac{1}{Q} \times 10^{\left(\frac{SENSI}{10} - 10.728\right)}\right) \times 100 \quad (14\text{-}28)$$

For 0.001 W, 30 ft:

$$\%\,Effici = \left(\frac{1}{Q} \times 10^{\left(\frac{SENSI}{10} - 5.978\right)}\right) \times 100 \quad (14\text{-}29)$$

For W/Pa/m:

$$\%\,Effici = \left(\frac{1}{Q} \times 10^{\frac{94\text{ dB} - \left[120 - 10.9 - 10\log\left(\frac{1\text{ W}}{\text{W/Pa}}\right)\right]}{10}}\right) \times 100 \quad (14\text{-}30)$$

Derivations

For 1.0 W, 1.0 m:

$$\frac{120 - 20\log\left(\frac{1\text{ m}}{0.282\text{ m}}\right)}{10} = 10.9 \quad (14\text{-}31)$$

For 1.0 W, 4 ft:

$$\frac{120 - 20\log\left(\frac{1.219\text{ m}}{0.282\text{ m}}\right)}{10} = 10.782 \quad (14\text{-}32)$$

For 0.001 W, 30 ft:

$$\frac{120 - 20\log\left(\frac{9.144\ \text{m}}{0.282\ \text{m}}\right) - 10\log\frac{1\ \text{W}}{0.001\ \text{W}}}{10} = 5.978 \tag{14-33}$$

For W/Pa/m:

$$\frac{1.0}{10^{\left(\frac{L_{EIA} - 44.78^*}{10}\right)} \tag{14-34}$$

$$*\ 94 + 10\log\left(\frac{0.001\ \text{W}}{1\ \text{W}}\right) + 20\log\left(\frac{1\ \text{m}}{9.144\ \text{m}}\right)$$

14.13.3 100% Efficiency Values

100% Effici = 109 dB at 1.0 W, 1.0 m, $Q = 1$.
100% Effici = 107.29 dB at 1.0 W, 4 ft, $Q = 1$.
100% Effici = 59.78 dB at 0.001 W, 30 ft, $Q = 1$.
100% Effici = 0.031 W for 1 Pa at 1.0 m, $Q = 1$.

14.14 Summary

Increased gain requires power. The complex combination of acoustic gain, loudspeaker sensitivity, efficiency, and required electrical power (manifested as a voltage at the loudspeaker backed by a sufficient current source) is a tapestry that properly woven yields dynamic powerful audio.

Bibliography

H. S. Antman. "Extension to the Theory of Howlback in Reverberant Rooms," *J. Acoust. Soc. Am.*, Vol. 39 (Feb. 1966).

W. K. Connor. "Experimental Investigation of Sound-System-Room Feedback," *J. Audio Eng. Soc.*, Vol. 21 (Jan. 1973).

D. Davis. "Equivalent Acoustic Distance," *J. Audio Eng. Soc.*, Vol. 21 (Oct. 1973), pp. 646-649.

V. V. Furduev. "Limiting Amplification of Sound in Closed Rooms," *Soviet Phys.-Acoust.*, Vol. 11 (Jan.-Mar. 1966).

William B. Snow. "Frequency Characteristics of a Sound Reinforcing System," *J. Acoust. Soc. Am.*, (April 1955).

W. Symmes. "Simplified Method of Calculating PAG/NAG Formulas Without Use of Logs," *Syn-Aud-Con Newsletter*, Vol. 1., No. 2 (1973).

R. V. Waterhouse. "Theory of Howlback in Reverberant Rooms," *J. Acoust. Soc. Am.*, Vol. 37 (May 1965).

Chapter 15

Designing for Speech Intelligibility

by Don Davis

15.1 Introduction . . . 267

The Talker or Performer . . . 267

The Microphone . . . 267

The Electronics System . . . 267

The Loudspeaker System . . . 267

The Acoustic Environment . . . 267

The Receiver-Listener . . . 267

15.2 Articulation Losses of Consonants in Speech . . . 268

Prediction, Measurements, and Anomalies . . . 268

Calculating $\%AL_{CONS}$ . . . 269

Usable Percentages . . . 269

15.3 Maxfield’s Equation . . . 270

15.4 Speech Power and Articulation . . . 270

15.5 Signal-to-Noise Ratio (*SNR*) . . . 271

15.6 Speech Intelligibility Calculations . . . 271

Calculating the Minimum Q . . . 271

Calculation of the Maximum RT_{60} . . . 271

Factors Affecting $\%AL_{CONS}$ . . . 272

Measuring Intelligibility . . . 272

Evaluating Speech Intelligibility Measurements . . . 273

Causes of Reduced Intelligibility . . . 274

15.7 Non-Acoustic Articulation Problems . . . 275

15.8 Relationship Between Q_{MIN} and $D_{2(MAX)}$ . . . 276

15.9 High Density Overhead Distribution . . . 276

15.10 $\%AL_{CONS}$ Variables . . . 277

The Role of Q in $\%AL_{CONS}$ . . . 277

The Role of $S\bar{a}$ and Ma in $\%AL_{CONS}$ . . . 277

Choosing the Correct $\%AL_{CONS}$ Equation . . . 277

Using a 1/3 Octave RTA to Obtain $\%AL_{CONS}$ . . . 278

Detecting the Presence of Detrimental Reflections Without Analyzers . . . 278

Relationship Between Acoustic Gain and $\%AL_{CONS}$ . . . 278

15.11 A Little History—Intelligibility Workshop 1986 . . . 278

The Results . . . 279

15.12 Summary . . . 279

Bibliography . . . 279

15.1 Introduction

The acoustic properties of rooms are basic to the successful design of sound systems. In this chapter we see that the ambient noise level in a room, the relationship of the direct sound to the reverberant sound level, and the reverberation time all affect the intelligibility of speech. Sound apparatus can be chosen and adjusted to optimize speech intelligibility.

The complete sound reinforcement system consists of:

1. Talker or other performer.
2. Microphone.
3. Electronics system.
4. Loudspeaker system.
5. Acoustic environment.
6. Receiver-listener.

15.1.1 The Talker or Performer

The performers are acoustic. They may use electro-acoustic adjuncts such as electronic instruments, guitars, and they may mix the input to the sound system by direct injection as well as acoustic inputs. The acoustic environment surrounding the microphone is unique unto itself and deserves separate analysis.

15.1.2 The Microphone

The microphone is an electro-acoustic transducer. Directional characteristics meaningful in free field can become meaningless in a reverberant sound field. A microphone's directivity is highly dependent on the acoustic environment. If it is used beyond D_c in a highly reverberant space it is essentially an omnidirectional device so far as its ability to discriminate against signals from loudspeakers, reflections, etc. When the signal, that is, the talker or some nearby performer, is within D_c, its directional characteristics come into play.

15.1.3 The Electronics System

Here correct initial sensitivity choices for microphones, careful gain structure and gain overlap, and the choice of adequate power to realize the potential acoustic gain desired can all be computed with great accuracy.

15.1.4 The Loudspeaker System

The loudspeaker system is another electro-acoustic transducer, which in many cases can act as both an input and an output device. The interaction of the driver and the applied coupling devices, horns, crossovers, etc., is a significant design choice.

15.1.5 The Acoustic Environment

It has been truly said that most sound system design ends about 4 ft in front of the loudspeaker. In actual fact, all sound system design should begin with the analysis of the acoustic environment in which the sound system is to operate. Noise levels, reverberation, delayed reflections, focusing, difficult geometry, and a myriad of other difficulties in the environment, must be recognized, analyzed, and corrected or adjusted to compromise in the system design.

15.1.6 The Receiver-Listener

Herein is the most challenging component. It can be analyzed but you have no direct control over it. Inarticulate talkers or handicapped listeners or some combination of both may be insurmountable and become the critics of your design.

Speech intelligibility centers on the 2 kHz octave band and so does the peaked response at the listener's eardrums, see Fig. 15-1. Note that at 2500 Hz the highest pinnae response is 72 dB and the lowest is 58 dB or a 14 dB difference. In tests that we conducted during an intelligibility workshop at Indiana University, we found that a person with a high pinnae response in the 2500 Hz region but with a poor frequency response, will score higher on an

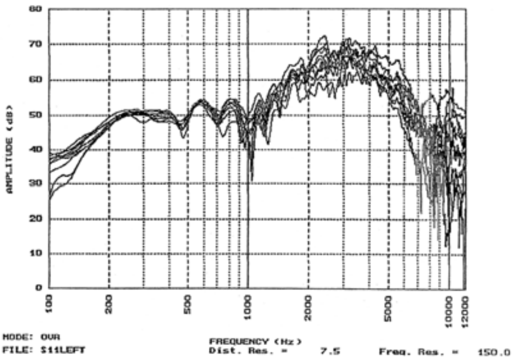


Figure 15-1. Composite of 30 left ears.

intelligibility word test than the person with a low pinnae response at 2500 Hz but with a near perfect frequency response curve. This information was gathered from 30 listeners that had very acceptable audiometric charts of their hearing. The pinnae response is largely due to the ear lobe configuration and their individual ear canal resonances. Hearing sensitivity, ear configurations that allow significant differences between ears, and training in focusing on speech can all provide very significant variations in speech intelligibility scores in individual listeners. For example, musicians will often ignore high harmonic distortion in a recording but immediately detect a wrong note or a pitch due to wow or flutter. Analysis of listeners is a psycho- acoustic task and awareness of listener types—ministers, choir directors, organists, allows "weighting" of their testimony.

15.2 Articulation Losses of Consonants in Speech

The articulation losses of consonants in speech ($\%AL_{CONS}$) are deeply interwoven into the total system design. They may be regarded as a separate acoustic entity only when the entire electroacoustic system is performing flawlessly. The widespread use of analysis promises increasingly analytical psychoacoustic research. In the meantime, the Peutz equations have usefully extended the work of the early pioneers of speech intelligibility to the point where their everyday use has proven their accuracy and practicality in sound system engineering.

15.2.1 Prediction, Measurements, and Anomalies

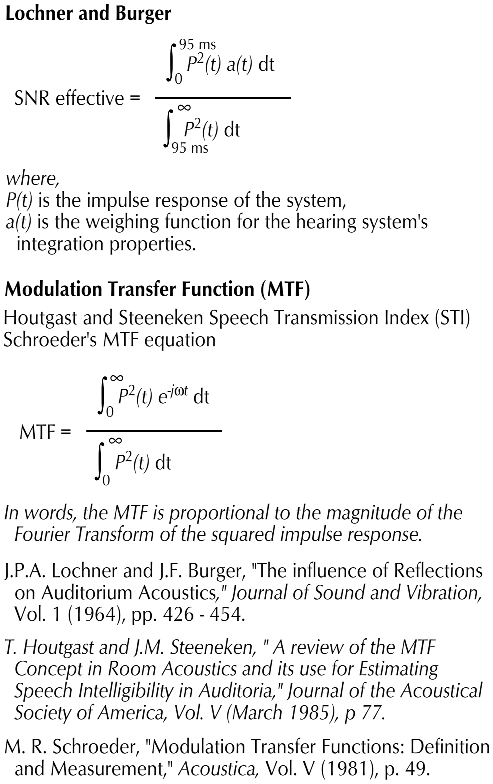
Sound system designers need to be able to predict speech intelligibility, measure it accurately, and understand the anomalies present in real environments that are not currently easily detected in advance of measurements.

Key parameters affecting speech intelligibility are:

1. Signal-to-noise ratio (*SNR*).
2. Reverberation time (RT_{60}), especially the level difference between the direct level and the reverberant level.
3. Distance from the source.
4. Source misalignments.
5. Reflections under 1 ft of path length difference.
6. Reflections that are late in time (100+ ms) and higher in level than energy near them.

The first three parameters are predictable within reasonable tolerances at the drawing-board stage. The last three are classified as anomalies that occur through oversight or error.

When the authors began their career in audio there was no practical way to predict speech intelligibility at the drawing-board stage of a sound system design. By 1953 Harvey Fletcher had written his remarkable book *Speech and Hearing in Communications* which did provide the means of measuring the intelligibility of a sound system through the use of articulation testing developed at the Bell Telephone Laboratories during the 1930s. Lochner and Burger provided insight into the role of signal-to-noise ratio (*SNR*) influence on speech intelligibility in a 1964 paper in *The Journal of Sound and Vibration*, Fig. 15-2.



Lochner and Burger

$$\text{SNR effective} = \frac{\int_0^{95\ \text{ms}} P^2(t)\ a(t)\ dt}{\int_{95\ \text{ms}}^{\infty} P^2(t)\ dt}$$

where,
P(t) is the impulse response of the system,
a(t) is the weighing function for the hearing system's integration properties.

Modulation Transfer Function (MTF)

Houtgast and Steeneken Speech Transmission Index (STI)
Schroeder's MTF equation

$$\text{MTF} = \frac{\int_0^{\infty} P^2(t)\ e^{-j\omega t}\ dt}{\int_0^{\infty} P^2(t)\ dt}$$

In words, the MTF is proportional to the magnitude of the Fourier Transform of the squared impulse response.

J.P.A. Lochner and J.F. Burger, "The influence of Reflections on Auditorium Acoustics," *Journal of Sound and Vibration,* Vol. 1 (1964), pp. 426 - 454.

T. Houtgast and J.M. Steeneken, " A review of the MTF Concept in Room Acoustics and its use for Estimating Speech Intelligibility in Auditoria," Journal of the Acoustical Society of America, Vol. V (March 1985), p 77.

M. R. Schroeder, "Modulation Transfer Functions: Definition and Measurement," *Acoustica,* Vol. V (1981), p. 49.

Figure 15-2. Signal-to-noise influence on speech intelligibility. (*Courtesy Journal of Sound and Vibration.*)

Meaningful prediction at the drawing board stage of articulation scores had to wait until 1971 when V. M. A. Peutz published his equation for the percent of articulation loss of consonants ($\%AL_{CONS}$) in speech.

Over a decade later, in the 1980s, Houtgast and Steeneken published their adaptation of the modulation transfer function (MTF) (first used in optics) to the speech transmission index (STI). During that same decade Peutz had evolved a most satisfactory measurement that computed the $\%AL_{CONS}$ from the measured L_D, L_R, L_N, and RT_{60}. At the present time only the Peutz equations allow a workable estimate of $\%AL_{CONS}$ at the drawing board stage, i.e., prior to any measurement. The MTF technique is a usable computation once actual measured data is obtained. As of the year 2005 there are computer programs that can model a room sufficiently to approximate the impulse response of the proposed room. See Fig. 15-2 for Lochner and Burger's as well as Mr. Schroeder's MTF integral defining the Houltgast-Steeneken technique.

For the reasons given here, we will concentrate on the V. M. A. Peutz equations and their techniques. They have the very real advantage of having software available that directly computes the $\%AL_{CONS}$ from Envelope Time Curve (ETC)* measurements. At this time, TEF analyzers are the most widely available analyzer in the hands of working sound contractors and acoustic consultants and allow, with existing programs, both the prediction and measurement of $\%AL_{CONS}$ with the same unit.

When we talk, the sounds we make can be broadly classified into consonants and vowels. The vowels are a, e, i, o, and u. Combinations like ba, pa, da, ta, ga, and tha contain consonant sounds. V. M. A. Peutz spent a number of years resolving that the percent of articulation loss of consonants determined the articulation score in various acoustical spaces. Formulas for $\%AL_{CONS}$, the articulation loss for consonants as a percentage, were then developed and published by V. M. A. Peutz and W. Klein of Nijmegen, Holland, in the December 1971 *Audio Engineering Journal*. The formulas have been adapted for this text by adding Q and presenting all their alternative forms. They can be useful in matching the room to the sound system.

15.2.2 Calculating $\%AL_{CONS}$

The formula for calculating the articulation loss of consonants as a percentage is:

$$\%AL_{CONS} = \frac{656 D_2^{\,2} RT_{60}^{\,2}(N)}{VQM} \qquad (15\text{-}1)$$

where,

D_2 is the distance from the loudspeaker to the farthest listener,

RT_{60} is the reverberation time in seconds,

V is the volume of the room in cubic feet,

Q is the directivity ratio,

N is the power ratio of L_W causing L_D to the L_W of all devices except those causing L_D,

M is the D_c modifier (usually 1 is chosen except in special instances).

The above formula is used for $D_2 \le D_L$, and $D_L = 3.16\, D_c$. When $D_2 \ge D_L$, the formula becomes:

$$\%AL_{CONS} = 9RT_{60} \qquad (15\text{-}2)$$

NOTE: It is necessary to assume a required *SNR* of 25 dB at the 2 kHz octave band to make these calculations valid. When meters are used for distances and volumes, the constant becomes 200.

15.2.3 Usable Percentages

Mr. Peutz states, "If the AL is below 10%, the intelligibility is very good. Between 10 and 15%, the intelligibility is good and only if the message is difficult and the speaker (talker) and/or listeners are not good will the intelligibility be insufficient. Above 15% the intelligibility is only sufficient for good listeners with speakers (talkers) and/or messages."

In comparing Peutz' method with known data from many installations, an AL_{CONS} of 15% is considered to be a practical working limit, yet many feel that one should strive for nothing more than 10%. The basic formula can be converted into the following useful forms:

$$\text{Max } D_2 \text{ for } 15\%\ AL_{CONS} = \sqrt{\frac{15VQM}{656\ RT_{60}^{\,2}(N)}} \qquad (15\text{-}3)$$

$$\text{Max } RT_{60} \text{ for } 15\%\ AL_{CONS} = \sqrt{\frac{15VQM}{656\ D_2^{\,2}(N)}} \qquad (15\text{-}4)$$

$$\text{Min } V \text{ for } 15\%\ AL_{CONS} = \frac{656 D_2^{\,2} RT_{60}^{\,2}(N)}{15QM} \qquad (15\text{-}5)$$

$$\text{Min } Q \text{ for } 15\%\ AL_{CONS} = \frac{656 D_2^{\,2} RT_{60}^{\,2}(N)}{15VM} \qquad (15\text{-}6)$$

---

*The ETC is related to a well-established concept in communication theory known as the modulation envelope. The ETC is the magnitude of the analytic signals description of the impulse response.

For a room with a volume of 250,000 ft³ (a medium-sized church) with a D_2 = 75 ft and for which a single horn with a Q = 20 has coverage angles that fit the audience areas, we could calculate the number of sabins we would like to have in the room by

$$\text{Max } RT_{60} = \sqrt{\frac{10\% \times V \times Q}{656 \times {D_2}^2}}$$

$$\text{Max } RT_{60} = \sqrt{\frac{10\% \times 250{,}000 \times 20}{656 \times 75^2}}$$

$$= 3.7 \text{ s}$$

and:

$$S\bar{a} = \frac{0.049V}{RT_{60}}$$

$$= \frac{0.049 \times 250{,}000}{3.7}$$

$$= 3311 \text{ sabins}$$

15.3 Maxfield's Equation

Peutz' formula was adapted to their data from one used by Western Electric's J. P. Maxfield for finding the "liveness" of a microphone pickup for broadcasting use in the late 1930s.

$$L = \frac{1000\ {RT_{60}}^2 {D_S}^2}{VQ} \tag{15-7}$$

where,

D_S is the distance from the talker to the microphone.

Acceptable values of L for speech range from 0.167 to 0.666. This equation can be manipulated in the same manner as the Peutz and Klein formula.

$$\text{Max } D_S = \frac{\sqrt{LVQ}}{31.6\ RT_{60}} \tag{15-8}$$

$$\text{Min } Q = \frac{1000\ {RT_{60}}^2 {D_S}^2}{VL} \tag{15-9}$$

$$\text{Min } V = \frac{1000\ {RT_{60}}^2 {D_S}^2}{LQ} \tag{15-10}$$

$$\text{Max } RT_{60} = \frac{\sqrt{LVQ}}{31.6\ D_S} \tag{15-11}$$

15.4 Speech Power and Articulation

There are two primary frequency response parameters of speech that require consideration by the sound system engineer. The first is the speech power as a function of frequency. Fig. 15-3 shows the typical distribution on a per cycle basis over the range 60–10,000 Hz for both men and women, and the ANSI curve (the third spectrum available from the GR 1382 RNG) is included for reference.

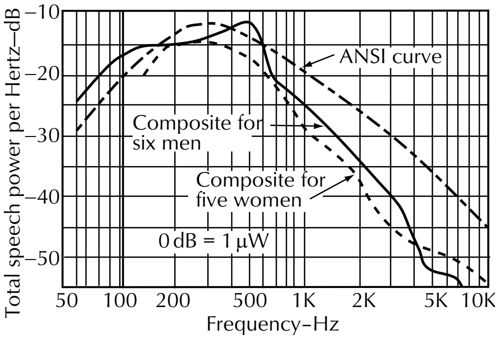


Figure 15-3. Relative speech power as a function of frequency for men and woman. (*Courtesy Journal of the Acoustical Society of America.*)

It is the spectrum shape that is important here as it gives an excellent idea of which frequencies will most likely receive the greatest power demands and the differences likely between one frequency and another in dB.

The second parameter is the relative contribution to intelligibility of each ⅓ octave band expressed as a percent of contribution to the articulation index, Fig. 15-4. Adding the percent contribution of each of the bands shown between 200 Hz and 4000 Hz equals 91.5% of the total contribution. The largest percentage, by far, is the 31.5% contribution in the 1.0 octave band centered on 2000 Hz. It can be seen why the telephone with its limited response works so well as do small radios with well-designed 4–8 inch loudspeakers covering the range of 125–5000 Hz.

It is of much greater importance to provide very smooth frequency response rather than extended frequency response and to control or eliminate specific nonlinearities in transducers that give rise to resonances, distortions, and other forms of coloration over the chosen response. When comparing a wide range high fidelity music system to a table-model radio, few realize how little difference the increased frequency response makes as compared to the differences in smoothness of

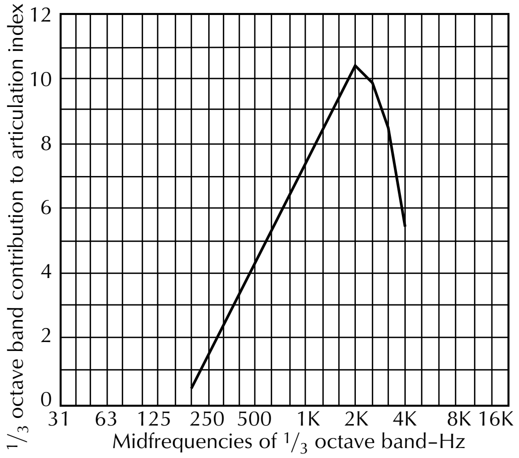


Figure 15-4. Variation of an articulation Index contribution with speech components in ⅓ octaves.

response through the very critical area from 250–5000 Hz.

Fig. 15-4 also confirms that if you can have the Q, RT_{60}, or sabins data at only one frequency, that one frequency should be the octave or ⅓ octave band centered on 2000 Hz.

15.5 Signal-to-Noise Ratio (*SNR*)

Fig. 15-5 illustrates the effect of signal-to-noise ratio, *SNR*, on the $\%AL_{CONS}$. The $\%AL_{CONS}$ improves steadily with improving *SNR* until it reaches 25 dB. After that, the articulation does not improve as the *SNR* is further extended.

This chart reveals that in rooms with an RT_{60} of 1.5 s, we would not want to accept an *SNR* of less than 25 dB. However, in a room of $RT_{60} = 0.5$ s, we could maintain a $\%AL_{CONS}$ of 15% with only an 11 dB *SNR*.

15.6 Speech Intelligibility Calculations

15.6.1 Calculating the Minimum Q

Given a room for which $V = 150{,}000\ \text{ft}^3$ and $RT_{60} = 1.92$ s, we can now calculate the minimum Q that will allow an AL_{CONS} of 15% at 125 ft in this room.

$$\text{Min } Q = \frac{656 D_2^{\,2} RT_{60}^{\,2}}{15V}$$

or

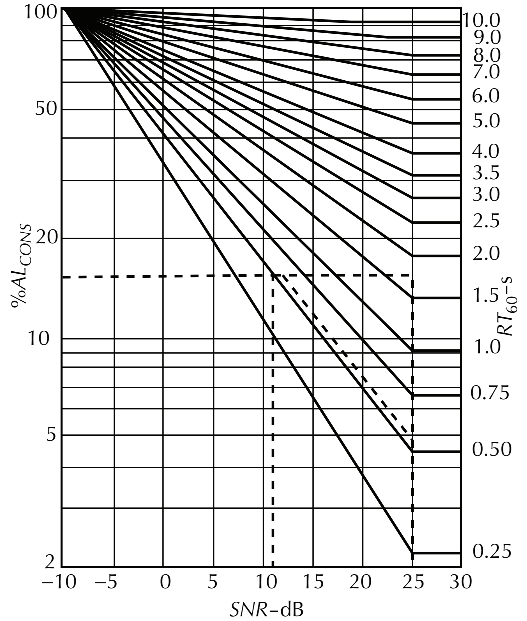


Figure 15-5. Effect of *SNR* on $\%AL_{CONS}$ at D_L.

$$\text{Min } Q = \frac{656 \times 125^2 \times 1.92^2}{15 \times 150{,}000}$$

$$= 16.8$$

Therefore, if you can locate a loudspeaker with proper angular coverage specifications that also has $Q \geq 16.8$, you can be sure of acceptable articulation at 125 ft.

There have been many striking confirmations of the accuracy of these formulae.

15.6.2 Calculation of the Maximum RT_{60}

Fig. 15-6 illustrates the $\%AL_{CONS}$ formulas in a simplified form. (The equations describe a curve, whereas the graph is a conservative straight-line approximation of the formula.) Referring again to Fig. 15-6, we see a baseline along the bottom expressing the source to listener distance in units of D_L. D_L is the limiting distance beyond which no further increase in articulation loss occurs. For a given acoustical environment, the limiting distance is related to the critical distance by

$$D_L = 3.16 D_C \qquad (15\text{-}12)$$

The left-hand vertical scale is $\%AL_{CONS}$ calibrated in percent, and the right-hand vertical scale is RT_{60} calibrated in seconds. Here we note that articulation losses continue to increase until D_L is reached.

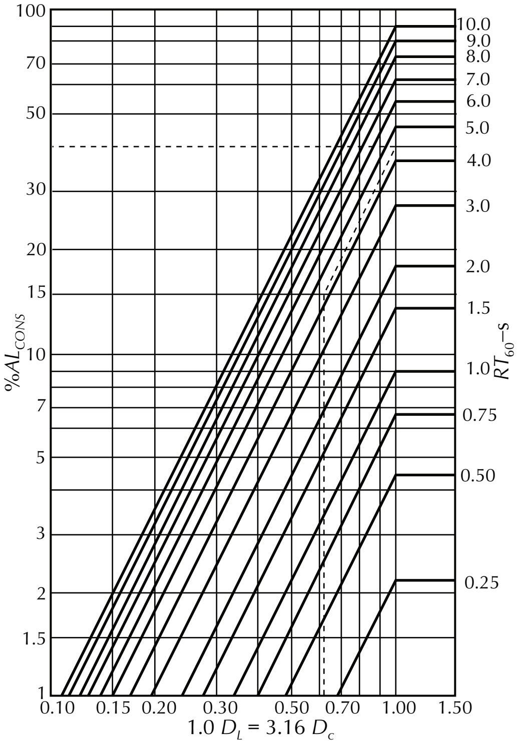


Figure 15-6. %AL_{CONS} as a function of source distance in units of D_L for various values of RT_{60}.

Beyond D_L, the articulation losses remain constant. Therefore, if we had a room with $RT_{60} \leq 1.6$ s and SNR of 25 dB, we could go any distance from the sound source without exceeding an AL_{CONS} of 15%.

Suppose, however, that we keep the SNR of 25 dB, but RT_{60} is now 4.5 s. We would follow downward between the slanting lines corresponding to 4 and 5 until we intersected the 15%AL_{CONS} horizontal line (or any other value of %AL_{CONS} being sought), and then drop straight down to the base line at 0.61 D_L. This would now be the new maximum distance from the source for which the articulation loss would be 15%.

Still another use of these very helpful equations is the determination of the limiting room parameters if an architect prefers a specific loudspeaker array for esthetic reasons. Say the architect has chosen a certain loudspeaker that has a $Q = 5$ and is designing a building having a maximum distance between the preferred loudspeaker location and the farthest listener of 150 ft. You can then use the equations to inform him of the maximum RT_{60} that will allow an AL_{CONS} of 15% at 150 ft from a loudspeaker with $Q = 5$. (Assume V is 150,000 ft^3.)

$$\text{Max } RT_{60} = \sqrt{\frac{15VQ}{656\ D_2^{\,2}}}$$

$$= \sqrt{\frac{15 \times 150{,}000 \times 5}{656 \times 150^2}}$$

$$= 0.873 \text{ s}$$

Using Eq. 15-12 for calculating D_L and if $S\bar{a}$ for this room is

$$S\bar{a} = \frac{0.049V}{RT_{60}}$$

$$= \frac{0.049(150{,}000)}{0.873}$$

$$= 8419 \text{ sabins}$$

$$D_L = 3.16 \times 0.141\sqrt{5 \times 8419}$$

$$= 91 \text{ ft}$$

The desired distance falls outside D_L, and the value of AL_{CONS} is 15% only because the limit 1.6 s was not exceeded.

15.6.3 Factors Affecting %AL_{CONS}

The %AL_{CONS} at a listener is dependent upon L_D, L_R, L_N, RT_{60}. It is also dependent upon a too-early or too-late return of reflected energy. A too-early return is usually under 3 ft and a too-late return is usually over 50 ft. The distance traveled or the signal delay is not the only factor. The amount of interference detrimental to %AL_{CONS} is also very "level" dependent. A too-early return that is visible on the analyzer, and audible to the ears will alter the amplitude response of the system and its polar response. A too-late return capable of interference will stand out above the normal exponentially decaying reverberation.

Much as the acoustic gain equations are made valid by the use of an equalizer when an amplitude aberration appears, so these too-early and too-late reflections are detected by measurements analysis and corrected by more careful placement of devices or selective absorption or diffusion.

The general case %AL_{CONS} equations work quite well provided no extraneous parameters such as the above go undetected at installation time.

15.6.4 Measuring Intelligibility

Prior to any measurements the area to be measured needs to be walked and listened to. Cup your ears to listen for focused reflections. Listen for high noise levels, lack of coverage, and any other anomaly. Intelligibility has to be investigated coverage area by coverage area.

Commissioning a sound system without full analysis that its speech intelligibility has been maximized is irresponsible. Audience areas should be carefully walked with speech information on the sound system (counting is very good), and questionable areas subsequently measured and corrected. Only then can you assuredly say this system is ready for use. The authors subscribe to the view that the most capable analyzer is still a trained pair of human ears, but "a trained pair of human ears" comes when ears have been calibrated by hours of listening coupled with using good measurement tools.

The trained listener has extraordinary capabilities not matched by any known measurement device. Take, for example, the oscilloscope, wave analyzer, etc., we can never be sure if its speech, music, or some randomly generated signal, but the connection of an inexpensive 2 inch loudspeaker will allow the listener to instantly determine which type of sound is present.

When a sound system has intelligibility problems in a given acoustic environment, the first test to make is to turn the sound system off and see if the same problems exist for a "live" talker. A live talker can be assumed to have a $Q = 2.5$ at the 2 kHz octave band with a coverage angle of 90° vertically by 120° horizontally for the same 2 kHz octave band.

Many mistake speech quality with speech intelligibility. It is important to separate judgment of sound quality from the issue of speech intelligibility. Speech intelligibility is defined by the score attained by live listeners to a live talker over the system in question. Sometimes intelligibility is sacrificed in order to present a wider bandwidth to the listener. Aircraft radios have high speech intelligibility but relatively strident sound quality. Speech over the telephone is intelligible but not necessarily wide range fidelity to the nuances of a given voice. Current measurement tools quantify only the speech intelligibility score. Quality, a subjective judgment, remains with the installer-user decision making.

When the design revolves around an existing building the time and expense of making measurements is very worthwhile. An ideal measurement technique is to take a portable hoist to the site, raise a loudspeaker of known directivity, and both talk and measure under and beyond critical distance, D_c.

A great deal of room analysis is done with impulse testing utilizing essentially omni-directional sources in order to maximize the excitement of the space. Since the goal of a successful sound system is the minimum excitation of the room, all measurements should use likely devices, on portable lifts, if necessary, to test from the proposed loudspeaker locations. In existing systems all tests should be conducted with the entire system in operation—one of the few times everything should be turned on at the same time when setting up the sound system.

Far too often measurements are made and corrections are undertaken without someone having stood in the presenter's position and talked unaided to the room. Often it is revealed that the room is unblemished and the deprecation of the sound quality and speech intelligibility over the sound system is due to the sound system itself. The authors have witnessed too many sound systems that raised the sound level but lowered the intelligibility.

In the past the only way to measure intelligibility was to use lengthy word lists and a group of live listeners. Elaborate testing procedures grew up around such processes and only on rare occasions was an effort made to actually check real life sound systems. Today we have very accurate and rapid methods of measuring intelligibility with acoustic instruments:

1. The Percent of Articulation Loss of Constants, $\%AL_{CONS}$.
2. The Speech Transmission Index, STI, Fig. 15-7.

15.6.5 Evaluating Speech Intelligibility Measurements

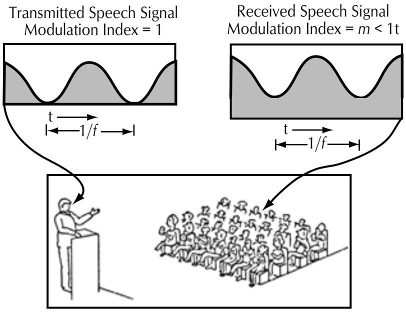
STI, Speech Transmission Index, measurements are a go-no-go-type analysis. They tell you there is a problem but do not contain any analysis of the cause or causes. If *SNR* happens to be the problem, STI works adequately.

The TEF analyzer's $\%Al_{CONS}$ measurements are more complex to make and to evaluate but they do contain the evidence necessary to allow the cause or causes to be isolated. Fig. 15-8 shows how to convert STI measurements to $\%Al_{CONS}$. Expecting close correlations between SI estimates and real life listeners is overly optimistic. Fine tuning of sound systems for maximum SI is done by ETCs, with careful attention to signal synchronization and equalization of the direct sound.

The ETCs obtained in this manner should, in a properly designed sound system, have no early reflections, under 10 ms. The proper place to divide direct from reflected sound is to measure only the direct sound. It is my opinion, though controversial, that early reflections under 30 ms of the direct sound do not aid intelligibility. They do raise the sound level perceived. To the best of my knowledge, it has

STI Intelligibility Measurement.

Speech Transmission Index, STI, is a method of quantifying the intelligibility of transmitted speech. Perfect transmission of speech implies that the temporal speech envelope at the listener's position replicates the speech envelope at the speaker's mouth. Speech intelligibility can be quantified in terms of the changes brought about in the modulation of the speech envelope as a result of noise and reverberation in the room.



The reduction in modulation can be described by a modulation reduction factor. The modulation reduction factor expressed as a function of modulation frequency is called the Modulation Transfer Function, MTF. This function provides an objective means of assessing the quality of the speech transmission, and from it, the STI value is derived. Both measurement methods are independent of statistical prediction as both measure in one form or another the modulation of the signal caused by whatever is helping or reducing the intelligibility.

Figure 15-7. STI intelligibility measurement. (*Courtesy B & K Instruments.*)

not been shown that the brain integrates early reflections except as loudness.

The Envelope Time Curve, ETC, is a very useful way to obtain data relative to speech intelligibility. Not only are ETCs reliable ways to gather L_N, L_D, L_R, and RT_{60}, but at the same time view the actual cause, if any, of poor intelligibility. Easily seen on the ETC screen is whether the causes are statistical, specific, or mixed.

As we have seen in our study of large-and small-room acoustics, the statistical basis behind the use of Sabinean-based equations is no longer valid when the reverberation time becomes small. Neither is prediction based upon those equations when used in such spaces. Fortunately, when the Sabinean-based equations are no longer valid, we are in an acoustical environment that cannot harm intelligibility through reverberation.

15.6.6 Causes of Reduced Intelligibility

1. Poor *SNR*.
2. Excessive reverberation (i.e., reduces $L_D - L_R$).
3. Specific long, delayed, high-level reflections.
4. Loudspeaker misalignment between alike devices.
5. Lack of synchronization (electronic).
6. Misequalization.
7. Poor quality devices (i.e., low Q when high Q is needed).
8. The distance from the source (i.e., increase $L_D - L_R$ by shortening D_2).

	STI	%AL_{CONS}		STI	%AL_{CONS}
	0.20	57.7		0.60	6.6
	0.22	51.8		0.62	5.9
BAD	0.24	46.5	GOOD	0.64	5.3
	0.26	41.7		0.66	4.8
	0.28	37.4		0.68	4.3
	0.30	33.6		0.70	3.8
	0.32	30.1		0.72	3.4
	0.34	27.0		0.74	3.1
POOR	0.36	24.2		0.76	2.8
	0.38	21.8		0.78	2.5
	0.40	19.5		0.80	2.2
	0.42	17.5		0.82	2.0
	0.44	15.7		0.84	1.8
	0.46	14.1		0.86	1.6
	0.48	12.7	EXCELLENT	0.88	1.4
	0.50	11.4		0.90	1.3
FAIR	0.52	10.2		0.92	1.2
	0.54	9.1		0.94	1.0
	0.56	8.2		0.96	0.9
	0.58	7.4		0.98	0.8
				1.00	0.0

$$AL_{CONS} = 170.5405 \times e^{(-5.419 \times STI)}$$

$$STI = [-0.1845 \times \ln(\%AL_{CONS})] + 0.9482$$

$$\text{i.e.: } [-0.1845 \times \ln(10.2)] + 0.9482 = 0.52$$

$$170.5405 \times e^{(-5.419 \times 0.52)} = 10.2\%$$

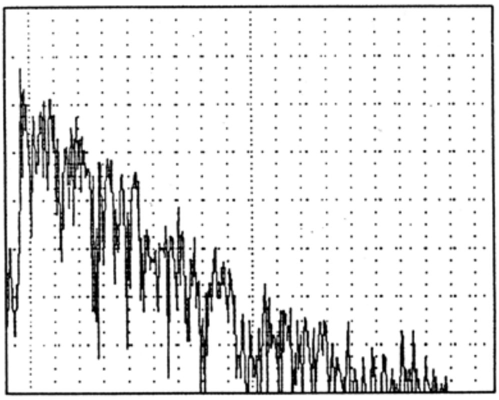
Figure 15-8. Converting STI measurements to %AL_{CONS}. (*Courtesy Farrel Becker.*)

Following are some of the methods to correct the problems causing reduced intelligibility.

***SNR* and Excessive RT_{60}.** The sound system engineer should have handled *SNR* by specifying a reasonable noise criteria (NC) curve as well as a suitable RT_{60} for octave bands 500Hz, 1kHz, and 2kHz.

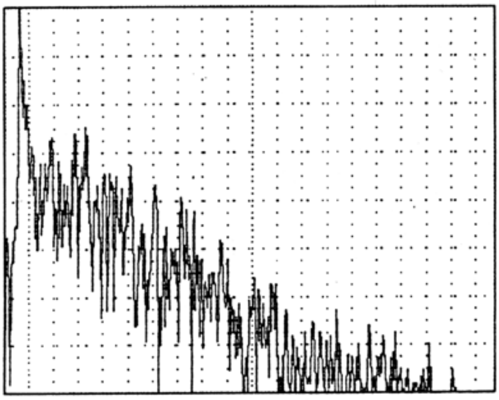
Specific Reflections. Correcting specific long, delayed, high-level reflections becomes a matter of loudspeaker placement and/or orientation. Only in rare cases is it a matter of room treatment.

Synchronization. Synchronization is usually used when electronic means are employed, usually a digital delay device, to synchronize two devices that must be physically separated but must be acoustically brought into identical arrival times at the listener's ears, Figs. 15-9, 15-10, and 15-11.



Vertical: 6 dB/division
Horizontal: 0–1,971,547 μsec
No delay

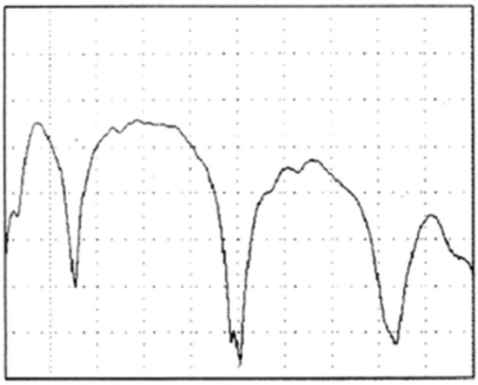
A. Between a far throw and near throw horn misaligned by 300 µs.



B. With a precision signal delay device correcting 300 µs misalignment (note increase in L_D.)

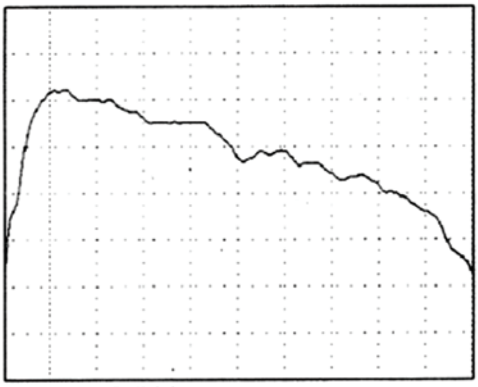
Figure 15-9. Show the before and after ETCs as measured in the audience area as a result of signal aligning (synchronizing) two horns. The measurements were made in the overlap area of the two devices.

Misalignment. Loudspeaker misalignment of alike devices is probably the most common cause of reduced intelligibility of sound systems. Misalignment causes spurious lobes to be radiated by the loudspeakers, which can excite wall surfaces. This causes an increase in the level of the reverberant sound field.



Vertical: 6 dB/division
Horizontal: 50.33–10,001.20 Hz

A. Two horns on at the same time.



Vertical: 6 dB/division
Horizontal: 50.33–10,001.20 Hz

B. Signal alignment used to delay the near throw horn 300 ms.

Figure 15-10. Shows the before and after frequency response in the overlap area.

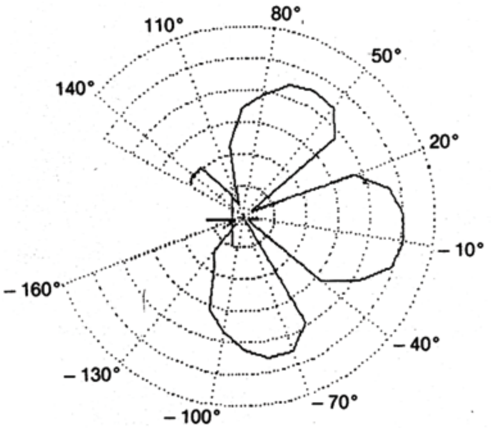
Misequalization. Misequalization also occurs more often than would be suspected. The missetting of levels often associated with the misuse of equalizers can result in intelligibility reduction because of premature distortion occurring ahead of the power amplifier.

Improper *Q* and Coverage Angles. There are loudspeakers in use today that have low intelligibility outdoors, to say nothing of their performance in a difficult acoustic environment. The most frequent misuse of devices is to use, in a reverberant space, a *Q* that is too low.

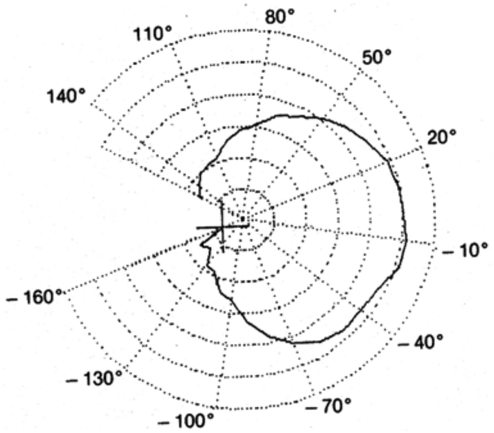
Distance from the Source. When all else fails, reducing the distance from the source to the listener will solve the problem. We have never been in an environment, where it's legal to work, that you could not communicate face to face.

15.7 Non-Acoustic Articulation Problems

Mispolarization of electronic components, misequalization of electronic components, or spurious



A. Out of physical alignment by 3 inches.



B. In physical alignment.

Figure 15-11. Shows the effect in the same audience area of the polar responses before and after alignment (synchronization).

oscillations due to miswiring or misgrounding can all lead to significantly reduced intelligibility.

The $\%Al_{CONS}$ equations presume that a properly designed, properly installed system is at hand and that the only remaining parameters to be considered are those in the acoustic environment. When there is poor intelligibility it is more often due to system problems than it is to acoustic environment problems.

15.8 Relationship Between Q_{MIN} and $D_{2(MAX)}$

It is not unusual to find that the $Q_{MIN(SS)}$ does not provide the necessary $C\angle$s required for even coverage of the audience area. If more than one device is to be tried at the original single source location (i.e., a "long throw" device of higher Q and a "short throw" device of lower Q), then the procedure is to increase $Q_{MIN(SS)}$ by N times.

$$Q_{MIN(long\ throw)} = N_{QMIN\,(SS)} \tag{15-13}$$

The Q of "short throw" horns is found by:

$$Q_{MIN(short\ throw)} = \frac{NQ_{MIN(SS)}}{4} \tag{15-14}$$

Dividing by a factor of four results in a short-throw device that has an on-axis level 6 dB lower than the long throw device for the same power input. It is important to note here that the L_R is determined by the power output of the devices and not by their sensitivity ratings. In a real reverberant sound field we want equal power contribution from each device. In a free field situation, you can vary the actual power to the device to obtain desired levels since your listener hears essentially only L_D. It can be disastrous to $\%AL_{CONS}$ to vary power inputs to devices in a truly reverberant space.

When $N_{QMIN(SS)}$ exceeds a realizable value, we resort to shortening D_2. When this option is called for, we usually choose a device with desirable coverage characteristics without paying undue attention to its Q. We then find N for the devices chosen by:

$$N = \frac{Q_{MIN(SS)}}{Q_{avail}} \tag{15-15}$$

Because N is increased and Q is not increased, we calculate a new $D_{2(MAX)}$

$$D_{2(MAX)} = \frac{D_{2(SS)}}{N} \tag{15-16}$$

This type of system has substituted a shorter D_2 for a required $NQ_{(SS)}$ and has maintained the D/R or $L_D - L_R$ required for the $\%AL_{CONS}$ involved. Signal delay is usually employed and the listener in the audience area receives a signal undistinguishable from a multiple device array at a single point. Failure to either increase Q or shorten D_2 when N is increased can dramatically affect $\%AL_{CONS}$.

15.9 High Density Overhead Distribution

Multiple sources generate "comb filter" amplitude responses. Interestingly, however, if enough sources are employed, the peaks of one comb filter response tend to fill the void of another comb filter response and our resultant curve smooths out. D_2 can, on rare occasions, become so short that the only answer becomes "pew back" loudspeaker systems (i.e., loudspeakers mounted in the back of pews or seats, usually with one loudspeaker for every two or three listeners).

15.10 %AL_{CONS} Variables

15.10.1 The Role of Q in %AL_{CONS}

Fig. 15-12 illustrates the effect for a loudspeaker with a $Q = 1$, a loudspeaker with a $Q = 5$, and a loudspeaker with a $Q = 50$. The effect of such dramatic changes in $L_D - L_R$ may or may not have dramatic effect on %AL_{CONS}. Once the reverberant sound field level has dropped below a certain point, it, like *SNR*, no longer really has any effect on the calculations or measurements at hand.

15.10.2 The Role of $S\bar{a}$ and Ma in %AL_{CONS}

When the $L_D - L_R$ is too negative a value, it can be altered by either changing L_D (i.e., increase Q or Me thus allowing L_W hence L_R to be lowered, or decrease D_2) or by changing L_R by lowering N or increasing Ma or $S\bar{a}$. Whenever given the opportunity to increase $S\bar{a}$ (assuming such an increase is needed), always do so first in a manner that directly increases Ma unless some special constraint such as a distant echo demands priority. Remember that in terms of %AL_{CONS} we prefer diffusion of the far reverberant sound field more than absorption as a means of altering its level; the Ma factor eliminates early reflections of potentially dangerous short intervals while at the same time reducing the acoustic power sent to the far reverberant sound field.

15.10.3 Choosing the Correct %AL_{CONS} Equation

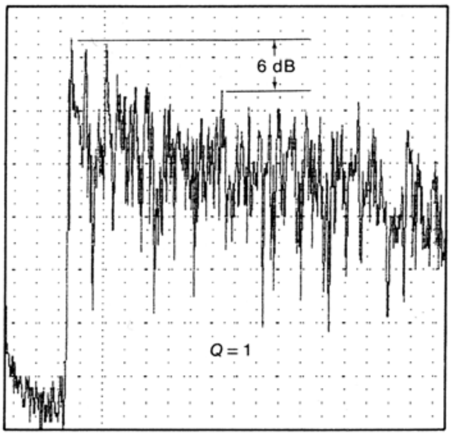
The Hopkins-Stryker equation provides an orderly analysis of L_W, Q, D_X, Me, N, $S\bar{a}$, and Ma and their contributions to L_D and L_R.

1. L_D is dependent upon L_W, Q, D_x, Me.
2. L_R is dependent upon $S\bar{a}$, Ma, N, L_W.

Therefore, we prefer the original Peutz equation for design purposes when at the drawing board stage as it keeps us in mind of the same key parameters:

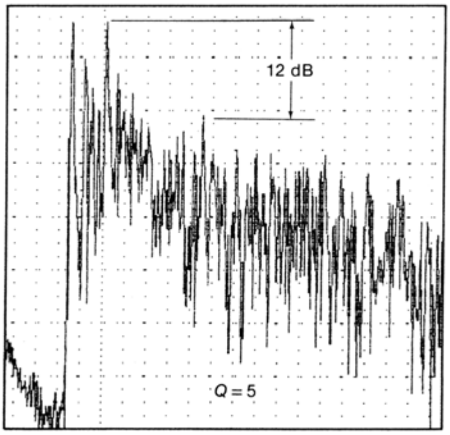
D_2, Q, Me, Ma, N, and RT_{60} (contains $S\bar{a}$).

If the building already exists or we are at the installation checkout stage, then the equations are our choice because their parameters are instantly accessible via analysis:



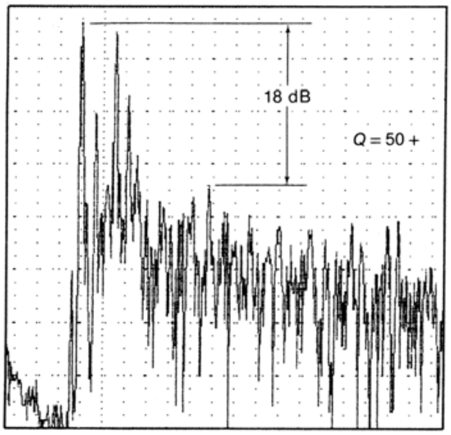
Vertical: 6 dB
Horizontal: 56,000-255,370 μsec

A. Q = 1.



Vertical: 6 dB
Horizontal: 56,000-255,370 μsec

B. Q = 5.



Vertical: 6 dB
Horizontal: 56,000-255,370 μsec

C. Q = 50.

Figure 15-12. The role of Q in %AL_{CONS}.

$$\frac{-0.32\log(L_R - L_N)}{10L_D + L_R + L_N} = A$$

$$-0.32\log\left(\frac{L_N}{10L_R + L_N}\log\frac{\frac{RT_{60}}{12}}{2}\right) = B \qquad (15\text{-}17)$$

$$10^{-2[(A+B)-AB]} + 0.015 = AL_{CONS}$$

where,

L_D is the direct sound level in dB expressed as a power ratio $L_D = 10^{(dB_D)/10}$,

L_R is the reverberant sound level in dB expressed as a power ratio $L_R = 10^{(dB_R)/10}$,

L_N is the ambient noise level in dB expressed as a power ratio $L_N = 10^{(dB_N)/10}$,

RT_{60} is the reverberation time in seconds for 60 dB of decay,

AL_CONS is the articulation loss of consonants expressed as a fraction.

15.10.4 Using a 1/3 Octave RTA to Obtain %AL$_{\mathrm{CONS}}$

L_D, L_R, L_N, and RT_{60} are also accessible through the use of conventional analyzers, albeit with some extra calculations. By measuring L_D near the source ("near the source" being defined as at least 10 dB above D_c), it can be extrapolated by inverse square law attenuation to the desired D_2. By measuring L_T at D_2 and L_N we can get an excellent idea of the L_R by:

$$L_R = 10\log\left(10^{L_T/10} - 10^{L_D/10} - 10^{L_{AMB}/10}\right) \qquad (15\text{-}18)$$

The RT_{60} is obtained in the usual manner.

15.10.5 Detecting the Presence of Detrimental Reflections Without Analyzers

If, in an otherwise acceptable environment, you find an audience area with unacceptable $\%AL_{CONS}$, a very useful tool is a panel of 4 inch "Sonex" with dimensions of approximately 4 ft × 4 ft. While someone "talks" the sound system, move the Sonex panel around the listener—above the listener, behind the listener, to the side of the listener, etc. Sometimes dual reflections are responsible and blocking at least two paths is advisable. Many first time users of this technique are startled at how dramatically $\%AL_{CONS}$ can improve as the Sonex panel intercepts the offending delayed signal that has too high a level for its arrival time.

One of the benefits of the recent practice of making the first three to four feet surrounding the loudspeaker grille as absorbent as possible is the elimination of many too early reflections (absorbent here means at the specular frequencies; such absorption is not active at the lower frequencies where, due to their wavelength, hard surfaces are not as detrimental).

15.10.6 Relationship Between Acoustic Gain and %AL$_{\mathrm{CONS}}$

Not only does the *SNR* constrain the maximum equivalent acoustic distance (*EAD*), but when there is failure to provide the necessary feedback stability margin (*FSM*) (6 dB as a minimum), then the sound system will be operated too near regeneration with the consequent result of nonlinear behavior of the peak energies. Coloration of this type is clearly audible and as a result $\%AL_{CONS}$ suffers.

15.11 A Little History—Intelligibility Workshop 1986

In the late summer of 1986 a major workshop on Speech Intelligibility was held under the auspices of Synergetic Audio Concepts with major participating manufacturers. Over 100 attendees participated in tests at three sites using three different sound systems at each site. One site was a reverberant Catholic cathedral, the second a national monument pre-sound film theater that had ideal speech acoustics due to its low noise, low reverberation environment. The third site was a new university concert hall with questionable geometry and severe speech intelligibility problems. At each site one sound system was low *Q* (1), a medium *Q* (approximately 7.0), and a high *Q* (over 25).

Three groups of thirty each took classic speech intelligibility tests in each of the three sites over each of the three systems. The intelligibility measurements were made using subjective word tests and Modified Rhyme Tests, MRT tapes by Dynastat. This was followed by TEF measurements and the B&K, then new RASTI instrument. All tests were monitored by three world renowned acoustical consultants, Rollins Brooks, David Klepper, and Victor Peutz plus the engineering personnel from the participating manufacturers.

15.11.1 The Results

1. Lobing that strikes multiple reflective surfaces in reverberant spaces is detrimental to intelligibility.
2. When well-behaved devices were used the original Peutz equation is remarkably accurate.
3. RT_{60} measurements should be of the Schroeder integration type.
4. High Q is only beneficial if aimed at listeners (absorptive). It is not beneficial aimed at walls, etc.
5. 10% is really about the maximum loss acceptable. Designing for 15% loss of constants sans audience usually results in 10% with audience.
6. The N factor is directly detrimental as it increases.
7. Delayed reflections at levels above that for other energy in the same zone are detrimental.
8. The signal-to-noise ratio, *SNR*, is vital to both prediction and measurement of speech intelligibility and should be at least 25 dB for the one octave band centered at 2 kHz.
9. Early returns, within 3 ft, at levels equal to or greater than the direct sound do affect speech intelligibility.
10. Misalignment and missynchronization measurably affect speech intelligibility.
11. High quality sound and high intelligibility are separate goals and not necessarily coincident.
12. In order to predict intelligibility scores you must have accurate, truthful, pertinent technical information. Q, directivity factor, is useless unless accompanied by truthful polar data for the region from 250 to 3000 Hz.

We wrote following the 1986 workshop that we sincerely doubted that a like-sized effort of this type would be made again and at this time, that is true.

15.12 Summary

You have become acquainted with calculating sound attenuation and level change under a variety of circumstances and have further found a way to calculate the effects of reverberation, directivity, and sound level on articulation losses at the listener's ears. TEF analysis, STI measurements, constant directivity loudspeakers, precision 10 µs signal delays, and an increasing understanding of the physical parameters underlying speech intelligibility in sound systems are leading the way to significantly improved speech reinforcement systems.

Bibliography

Don Davis and Carolyn Davis. "Application of Speech Intelligibility to Sound Reinforcement," *J. Audio Eng. Soc.*, Vol. 37, No. 12 (Dec. 1989).

H. Fletcher. *Speech and Hearing in Communication*. New York: Van Nostrand, 1953.

W. Klein. "Articulation Loss of Consonants as a Basis for the Design and Judgement of Sound Reinforcement Systems," *J. Audio Eng. Soc.,* Vol. 19 (Dec. 1971).

V. M. A. Peutz. "Speech Reception and Information," *Syn-Aud-Con Tech Topic*, Vol. 5, No. 12 (1978).

__________. "Articulation Loss of Consonants as a Criterion for Speech Transmission in a Room," *J. Audio Eng. Soc.*, Vol. 19 (Dec. 1971).

Chapter 16

What is Waving and Why

by Eugene Patronis, Jr.

16.1 General Properties of Air 283
16.2 Plane Waves 288
16.3 Non-Planar Wave Motion in a Tube 297
16.4 Plane Wave Tubes having Arbitrary Terminations 299
16.5 Impedance Tube 303
16.6 More General Waves 305
    Pulsating Sphere 305
16.7 Acoustic Intensity 308
16.8 Boundaries 309
16.9 Acoustic Dipole 310
Bibliography 311

Unlike electromagnetic waves that can exist in a vacuum as well as in material substances, sound waves are mechanical waves and require a material medium in which to exist. The medium may be either a solid such as a bar of steel or a fluid such as water or air. Fluids are distinguished from solids in that a fluid will assume the shape of the container in which it is placed. If the fluid in question is a liquid of small volume, however, it will not occupy all of the space provided by a container of larger interior volume whereas a gaseous fluid will occupy all of the interior space provided by the container. In the process of doing so, the pressure, temperature, and the energy content of the gas must undergo adjustments to make this possible. The objective here is to learn in a fundamental way that wave properties such as speed, dispersion, momentum transfer, energy transport, and guidance depend upon both physical law and the properties of the host medium. Along the way we will encounter all of the familiar wave properties such as interference, diffraction, and refraction as well as phase and group velocity. It is a fact that air borne sound plays a dominant role in sound reinforcement and room acoustics so we will begin our study with air considered to be the supporting medium for our treatment of sound waves.

16.1 General Properties of Air

Air is a mixture of several different gaseous components with the principal ones being displayed in Table 16-1 as molecular fractions of the total composition. The numbers displayed as decimal fractions are for dry air. Normal air also contains water vapor. This does not appear in the table, as it is a varying quantity depending upon the weather conditions. The major acoustic influence of the moisture content of air is that of a frequency dependent conversion of sound energy into heat. This process will be described at the appropriate time.

Table 16-1. Principal components of dry air at sea level

Gas	Symbol	Fraction
Nitrogen	N_2	0.78040
Oxygen	O_2	0.20946
Argon	Ar	0.00934
Carbon dioxide	CO_2	0.00038

Note that if you sum the fraction column the result will be slightly less than one. The reason for this is that dry air also contains trace amounts of hydrogen, helium, neon, krypton, xenon, radon, and methane. The fractions presented in the table are the molecular fractions rather than the mass fractions. This means that if you take a sample of dry air at sea level 78.040% of the molecules in the sample will be diatomic molecules of nitrogen while 20.946% will be diatomic molecules of oxygen, etc. We could construct a similar table where fractions of the total mass rather than fractions of the total number of molecules represent the various components. The mass fractions would be different because, for example, molecules of oxygen have greater mass than do molecules of nitrogen. In discussing molecular masses associated with a sample of substance that you might encounter in everyday life, i. e., a macroscopic sample we usually refer to a mole of the substance. The mole is one of the seven SI base units and is called quantity of substance. The mole has the value $6.02(10^{23})$ which is Avogadro's number. You can have a mole of anything. A mole of dollars would be $\$6.02(10^{23})$. With that amount of money you could give to each person on earth an amount equal to the national debt of the United States and never miss it! A mole of dry air has a mass of 0.02898 kilogram (kg). This is called the molar mass of dry air. The molar mass will be represented in this article by M.

Reference conditions for a standard atmosphere are usually taken as the pressure at sea level with a temperature of 0° C. This temperature on the Celsius or centigrade scale corresponds to 273.15 K on the absolute or Kelvin temperature scale. You should note that we did not write °K because K alone stands for degrees on the absolute temperature scale. Note also that the degree increments on each scale are the same size so any temperature reading on the Celsius scale can be converted to absolute simply by adding 273.15. The standard atmosphere has a static or undisturbed pressure of $1.01325(10^5)$ Pascals (Pa) and a density of $1.293\,\mathrm{kg{\bullet}m^{-3}}$. A sound with a *SPL* of 94 dB corresponds to an rms acoustic pressure of 1 Pa. This moderately loud sound, which represents a disturbance away from static conditions, perturbs the atmospheric pressure less than one part in 10^5. Accompanying the pressure disturbance there are also disturbances in the local air's density and temperature. In order to understand this we must consult what is called the equation of state of a gas. Furthermore the sound source feeds acoustic energy into the air. In order to get a handle on this we must consult the first law of thermodynamics.

The equation of state of an ideal gas as given by the ideal gas law says

$$PV = nRT \tag{16-1}$$

where,

P is the total pressure exerted by the gas on the walls of the containing vessel,

V is the interior volume of the container,

T is the absolute temperature of the gas,

n is the number of moles of the gas in the container,
R is the universal gas constant.

The value of R is 8.3145 J•mole^{-1}•K^{-1}. Joule is abbreviated J. An ideal gas would be one in which collisions between individual molecules of the gas as well as collisions with the walls of the container would be perfectly elastic, in which both momentum and kinetic energy are conserved in the collision processes. Real gases do not obey the ideal gas law under all possible conditions. The ideal gas law was determined by experimentally studying the behavior of real gases as a function of the gas density. With sufficiently low densities all real gases were found to follow the same equation of state that has become to be known as the ideal gas law. The behavior of air in the temperature and pressure ranges that we normally encounter in everyday life follows the ideal gas law with little error. We should note at this point that the ideal gas law could be expressed by using the gas density rather than the gas volume as a variable of interest. Let the total mass of our sample of gas be represented by m. Then the number of moles in our sample of gas would become m/M and we can write the following sequence of equations:

$$\begin{aligned} PV &= nRT \\ &= \frac{m}{M}RT \end{aligned} \qquad (16\text{-}2)$$

$$\begin{aligned} P &= \frac{m}{V}\frac{R}{M}T \\ &= \rho\frac{R}{M}T \end{aligned} \qquad (16\text{-}3)$$

In the last equation the mass per unit volume or density of the gas is represented by the Greek letter rho (ρ).

The first law of thermodynamics in simple terms states that the change in the internal energy of a physical system is equal to the heat energy added to the system less the work done by the system. If we let Q represent the added heat energy and W represent the work done, then $\Delta U = Q - W$, where ΔU stands for the change in internal energy. In the case of a system composed of an ideal gas, the internal energy is associated only with the kinetic energies of the molecules composing the gas. The individual molecules in such a gas undergo random motions throughout the volume of the gas and have speeds that can change from moment to moment as a result of collisions with other molecules or with the walls of the containing vessel. Monatomic molecules such as Ar can only have kinetic energies associated with translations in the three perpendicular spatial directions. We say such molecules have three degrees of freedom. Diatomic molecules such as N_2 and O_2 in addition to translation can have kinetic energies associated with rotations about two perpendicular axes. Such molecules are said to have five degrees of freedom. Finally, polyatomic molecules such as CO_2 can potentially perform distinct rotations about three mutually perpendicular axes and have six degrees of freedom. Kinetic theory tells us that for each molecule in the gas there is, on the average, a kinetic energy that is proportional to kT where k is Boltzman's constant with k being equal to R divided by Avogadro's number or $1.38(10^{-23})$ J•K^{-1}. The significance of this is that the internal energy of an ideal gas or a real gas that behaves as an ideal gas is directly proportional to the absolute temperature alone. Regardless of the complexity of a molecule's structure, each molecule in the gas has an average translational kinetic energy of $3/2kT$. Knowledge of a molecule's average kinetic energy allows the calculation of the root mean square molecular speed. The formula for this calculation is

$$v_{rms} = \sqrt{\frac{3RT}{M}} \qquad (16\text{-}4)$$

If you apply this formula to air at standard conditions v_{rms} will be found to be 485 m•s^{-1}.

As a result of many experiments it has been determined there are two types of air compression and expansion processes associated with sound waves. The first and simplest of these occurs normally at ultrasonic frequencies, well above the audible spectrum but can also occur in a sealed box loudspeaker enclosure which is completely filled (except for the volume occupied by the loudspeaker) with loosely packed fiberglass. This is the isothermal process. The second process is known as the adiabatic process and is applicable in free air throughout the audible spectrum and beyond. An isothermal process is one in which the gas temperature and hence its internal energy remains constant. An inspection of the ideal gas law for a fixed quantity of gas and a fixed temperature reveals that the isothermal process is described by $PV = C_T$ where C_T is just a constant. For one mole of air at standard conditions $C_T = 2270$ J. An adiabatic process is one in which heat energy is neither added nor subtracted. For an adiabatic process not only must the pressure-volume relationship follow the ideal gas law it must also satisfy $PV^{\gamma} = C_Q$ where C_Q is a constant that depends on the state of the gas and γ is a constant for the particular mixture composing the gas. For air γ has the value of 1.402 and is dimensionless whereas for one mole of air at standard conditions C_Q has the value 493 with the strange dimensions of $Pa \bullet (m^3)^{\gamma}$. A mole of any ideal gas under the standard conditions of sea level atmospheric pressure and a temperature of 273.15 K occupies a volume of 0.0224 m³ regardless of the type of expansion or compression process that may occur. This is illustrated in Fig. 16-1 for one mole of air where a comparison is made between plots of the pressure-volume relationship for both an isothermal and an adiabatic process.

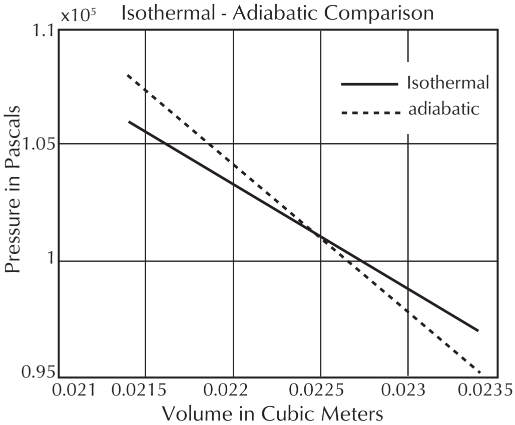


Figure 16-1. Isothermal-Adiabatic comparison. The two curves have a common point at standard pressure and temperature with a volume of 0.0224 m³.

At the intersection point of the two curves, the gas pressures, volumes, and temperatures are the same. In the isothermal process, the first law of thermodynamics tells us that when air does an amount of work W against an outside agency in expanding beyond the intersection point, an amount of heat Q equal to W must be added to the air in order to maintain the temperature T at a fixed value. Conversely when an outside agency compresses the gas and thus does work on the gas while reducing its volume, a corresponding amount of heat energy must be removed from the gas in order to maintain a constant T. With a constant temperature there is no change in the internal energy of the gas in either circumstance. In the adiabatic process, however, no heat energy is added or removed so when the gas does work during expansion, the gas itself must supply this energy. As a consequence the internal energy of the gas decreases and the gas temperature drops. Conversely, of course, when an outside agency compresses the gas adiabatically, work is done on the gas rather than by the gas and the internal energy of the gas increases by an amount equal to the work done. This manifests itself as an increase in gas temperature.

We have mentioned the work done on or by a gas several times without divulging how the work is determined. We hasten now to remove that omission. Suppose the gas is air and it is contained in a cylinder that has a tightly fitted piston upon which we can exert a sufficient force to move the piston inward so as to compress the gas adiabatically. Alternatively, we can relax our force on the piston somewhat and allow the gas to expand adiabatically and thus push the piston outward. In order to do this the entire apparatus must be thermally insulated from the external environment so that heat energy cannot leak in or out during the process. The formal definition of the work done by the gas in the process of the volume changing from an initial value V_i to a final value V_f is given by the following integral equation.

$$W = \int_{V_i}^{V_f} P dV \tag{16-5}$$

This integral has a geometrical interpretation. It actually calculates the area that lies beneath the plot of pressure versus volume between the limits of V_i and V_f. When the final value of volume is greater than the initial value the numerical value of the calculation is positive indicating that the gas has done work against the piston while expanding. If the opposite is true, final volume less than initial volume, the numerical value is negative indicating that the piston has done work on the gas by compressing it. We should also observe that if we multiply the dimensions of pressure by the dimensions of volume we obtain the dimensions of energy thusly, Newton•m^{-2}•m^3 = Newton•m. A Newton•m is of course a Joule. An example of this is displayed in Fig. 16-2 where a mole of air is compressed from an initial volume of 0.022 m³ to a smaller volume of 0.018 m³.

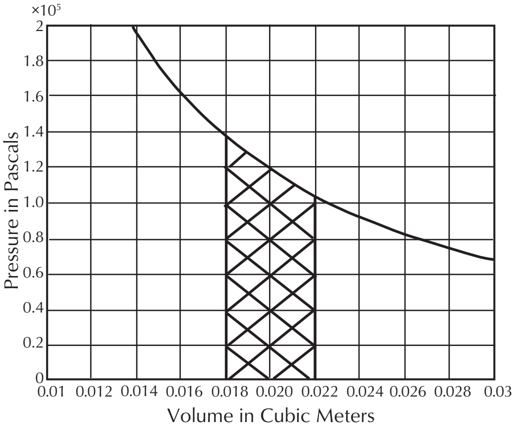


Figure 16-2. Adiabatic compression of air. The cross-hatched area represents the work done on the gas during compression.

The value of the integral is –478 J with the minus sign indicating that work was done on the gas. In other words, the agency pushing the piston did a work of 478 J in compressing the gas. This amount of energy is potential energy stored in the compressed gas because the gas could do this amount of work on the piston in expanding back to its original state. You can visually do a rough calculation of the area under the graph by noting that each complete crosshatched block has an area of 0.002•20,000 or 40 J and there are 11 complete blocks and a slightly less than complete partial block. The eleven complete blocks correspond to 440 J.

Finally, we are ready for an acoustical calculation of some importance. We previously mentioned that an adiabatic process was described by $PV^{\gamma} = C_Q$ where the constant C_Q depends on the quantity of air involved. An alternative description of an adiabatic process for air in the atmosphere that is independent of the amount of air involved is

$$P = \left(\frac{\rho}{\rho_0}\right)^{\gamma} P_0 \tag{16-6}$$

In this equation P is the total air pressure when the air density is ρ and P_0, ρ_0 are the values of the air pressure and density under standard conditions. Fig. 16-3 illustrates this behavior.

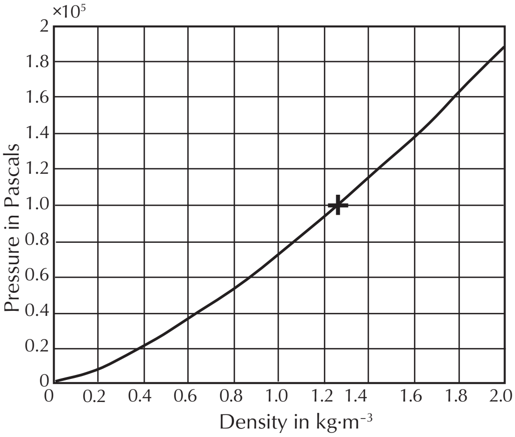


Figure 16-3. Air pressure versus density. The marker indicates the location of the point (P_0, ρ_0).

If we examined the curve in the vicinity of the marker on a greatly magnified scale the curve would appear to be a straight line and we could determine its slope graphically. This slope describes the ratio of a small change in pressure to a small change in density. Alternatively, we could employ the methods of differential calculus and precisely determine a value for the slope. Calling this slope c^2, the slope of the curve at the position of the marker is

$$c^2 = \gamma\frac{P_0}{\rho_0} \tag{16-7}$$

Upon inserting the values for standard air pressure and density c^2 is found to be 109866.71 with the dimensions of Newton•m•kg^{-1} = m^2•s^{-2}. The obvious next step is to extract the square root to find $c = 331.46$ m•s^{-1}. This result should be familiar as it is the speed of sound in air at a temperature of 273.15 K! This of course is interesting and the explanation for it will be forthcoming. It is not, however, the calculation that we seek at this point. Let us turn our attention to a cubic centimeter of air under standard conditions. A cubic centimeter (cm^3) is 10^{-6} cubic meter (10^{-6} m^3) so our sample contains a mass of $1.293(10^{-6})$ kg of air. Let our air sample be in contact with some vibrating object such that the sample is momentarily compressed by a very small amount. The sample of air now occupies a smaller volume so its density has increased and this is accompanied by an increase in the air pressure in the sample itself. Let the new pressure be $P = 101326$ Pa. The undisturbed pressure was $P_0 = 101325$ Pa. What we call acoustic pressure, p, is the difference between these two numbers so $p = P - P_0 = 1$ Pa. Now we can use the slope on Fig. 16-3 to calculate the change in density from $p = c^2(\rho - \rho_0)$. Remember the slope is pressure change divided by density change. We can solve for ρ to find $\rho = 1.293009102$ kg•m^{-3}. We started with a mass of air of $1.293(10^{-6})$ kg that was contained in an initial volume V_0 of 10^{-6} m^3. We still have the same mass of air but now in a smaller volume V so we can write $\rho_0 V_0 = \rho V$. Solving for V we find that $V = 9.99929606(10^{-7})\ \mathrm{m}^3$. The change in volume in the compression process is then $V - V_0 = -7.039393602(10^{-12})\ \mathrm{m}^3$. The result of this is that there is now energy stored in our compressed sample of air. A portion of this stored energy is called acoustical potential energy and is represented by the work that can be performed by the acoustic pressure as the sample expands back to its original volume against the external pressure of the surrounding air. This work can be determined from Fig.16-4 in which the acoustic pressure is plotted versus the size of the volume change.

The work that can be performed by the acoustic pressure as the sample expands back to normal size is the area included in the triangle and in this case is one-half the altitude times the base or $0.5•7.039393602(10^{-12})$ Pa•m^3 = $3.5197(10^{-12})$ J. We can calculate this directly using the equation for acoustical potential energy that is obtained through the full use of calculus.

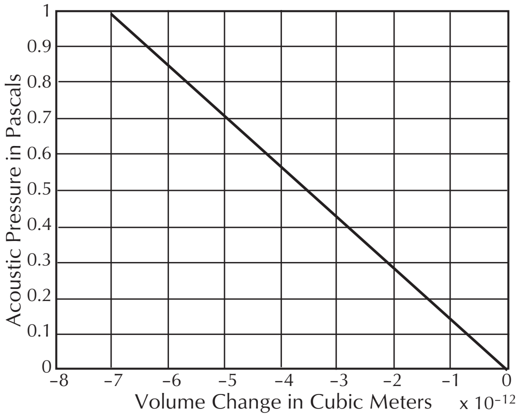


Figure 16-4. The acoustical potential energy is the area of the triangle.

$$E_P = \frac{1}{2}\left(\frac{p^2}{\rho_0 c^2}\right) V_0 \qquad (16\text{-}8)$$

The equation given above is essential to the calculation of a part of the energy transported by a sound wave.

You will recall that the acoustic pressure is given by $p = P - P_0$. In this equation, p is the acoustic pressure at some point in space and some instant in time. Similarly, P is the disturbed total atmospheric pressure at the same point in space and the same instant in time while P_0 is the static or undisturbed atmospheric pressure at the location of interest.

The acoustic pressure is perhaps the premier acoustic variable. The root mean square value of the acoustic pressure at a particular location expressed in Pascals is what is used in determining the sound pressure level at that location through the relationship

$$SPL = 20 \text{ dB} \log\left(\frac{p_{rms}}{2(10^{-5})}\right) \qquad (16\text{-}9)$$

There are, however, many other important acoustic variables whose values depend upon location in both space and in time. A listing of the ones to be employed here appears in Table 16-2.

Table 16-2. A Partial Listing of Acoustic Variables

Name of Variable	Symbol	Unit
Acoustic Pressure	p	Pa
Air Density	ρ	$\text{kg}\bullet\text{m}^{-3}$
Condensation	s	dimensionless
Particle Displacement	ξ	m
Particle Velocity	u	$\text{m}\bullet\text{s}^{-1}$

Table 16-2 introduces three acoustic variables that we have not yet discussed. The acoustic condensation symbolized by s is simply the ratio of the change in air density brought about by an acoustic disturbance to the normal static or undisturbed air density as expressed by the equation

$$s = \frac{\rho - \rho_0}{\rho_0} \qquad (16\text{-}10)$$

In this equation the Greek letter rho, ρ , represents the total density of air under the disturbed condition while ρ_0 represents the undisturbed or static air density. The acoustic variable that is termed the particle displacement and is symbolized by the Greek letter xi, ξ, will require a more lengthy explanation. The question that immediately arises is what constitutes an air particle? It cannot be a single molecule as air is always composed of a collection of a variety of molecules in the proportions tabulated in the beginning of this chapter. The particle size, whatever its value, must be sufficiently large so as to encompass millions of molecules in order to yield valid statistical averages and thus behave as an apparently continuous fluid while at the same time it must be small enough that the acoustic variables are essentially constant throughout the volume occupied by the particle. This latter condition requires the dimensions of the particle to be very much smaller than any sound wavelength under consideration. Let's do a simple calculation in order to determine a reasonable size for what we will call an air particle. What volume would say two million molecules of air occupy under standard conditions? We learned earlier on that Avogadro's number of molecules would occupy about 0.0224 m^3 under standard conditions. If we consider a cube of edge dimension l then by simple proportion we can write

$$\frac{l^3}{2(10^6)} = \frac{0.0224}{6.02(10^{23})}$$

When this is solved for l the result is found to be $4.2(10^{-7})$ m. This distance is orders of magnitude smaller than the wavelengths encountered in air even at ultrasonic frequencies so both of our requirements are satisfied. We might even round this number upward to a value easier to remember and say that an air particle is that amount of air under standard conditions that occupies a cube having an edge dimension of about 0.5 micron. When we consider such a small cubical volume of air that we now will call an air particle we realize that even in the absence of an acoustical disturbance, the air molecules are constantly undergoing random thermal motion. As a result of this thermal motion, some molecules move out of the volume but other molecules having the same properties also move into the volume. The volume has been chosen large enough so that the randomness of the thermal motion averages to zero so that in effect the air contained in the particle is at rest. An acoustical disturbance, as we shall see, imposes a preferred direction of motion and thus can bring about a displacement of the air particle as a whole. Particle displacement is a vector quantity and as such has both a magnitude and a direction that are measured relative to a coordinate frame of reference. In addition to particle displacement we will also be concerned with another acoustic variable that describes the instantaneous rate at which the particle displacement changes with time. This is a vector quantity also and is called the particle velocity. As the table indicates the symbol employed for the particle velocity is the letter u.

In the absence of any acoustical disturbance the acoustic variables of Table 16-2 are all zero with the exception of ρ for which the value becomes the static atmosphere value ρ_0. When an acoustical disturbance is present all of the acoustic variables listed in the table will have values that depend upon both location in space and time. For simplicity let's center our attention on just the acoustic pressure as an example. Mathematically we say that the acoustic pressure is a function of the positional coordinates and time. If we were employing general Cartesian coordinates this mathematical statement would be written in the manner, $p = p(x, y, z, t)$.

The entry immediately above is read as, "The acoustic pressure is a function of x, y, z, and t." It does not tell you what particular mathematical function but only that there is such a function. In certain situations not all spatial coordinates may be involved. If the acoustic pressure depends only on the z coordinate and time then $p = p(z,t)$ would be appropriate. From either theory or experiment we may find what the particular mathematical functional dependence is. For example, the answer might be $p = p_m \cos(\omega t - kz)$.

In this answer p_m, ω, and k are constants and we are informed that the acoustic pressure varies as the cosine of the difference of two angles one of which is directly proportional to time and the other of which is directly proportional to the value of the z coordinate. As we shall see shortly this function describes a plane wave propagating in the direction of increasing values of the z coordinate.

16.2 Plane Waves

Now that we have covered the preliminaries, we turn our attention to a physical system of some importance consisting of a long, rigid-wall air filled pipe. The inner diameter of the pipe is d, its inner radius is a. The interior wall is smooth and the pipe is straight. The wall thickness of the pipe is immaterial as long as it is reasonably rigid. We will employ cylindrical coordinates for locating positions in the pipe. These are the coordinates best suited for such a structure. In cylindrical coordinates space points are located by the variables r, θ, and z. The z coordinate is familiar from the usual Cartesian set. The relationship between r, θ and the familiar x, y can be extracted by viewing Fig. 16-5.

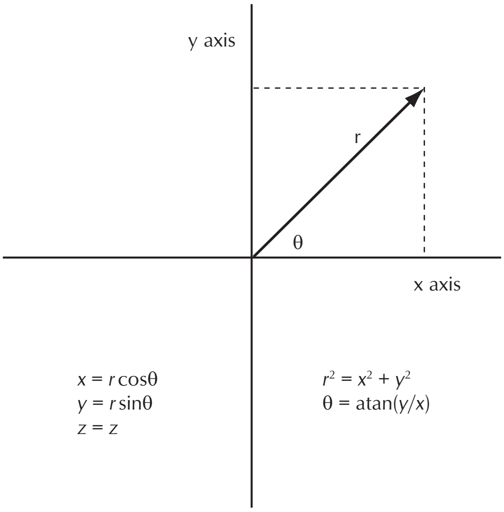


Figure 16-5. Cartesian to cylindrical conversion. The z axis points toward the reader.

We have selected an air filled pipe as the starting point for our discussion of acoustic waves because of the ease with which the simplest of wave motions, namely plane waves, can be established in such a structure. Our first step will be to concentrate on a small mass of air in the pipe under static conditions and on the same mass of air after it has been acoustically disturbed. The physical situation is depicted in Fig. 16-6.

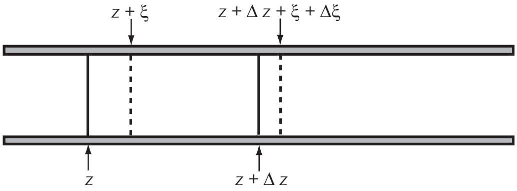


Figure 16-6. Undisturbed (solid) and disturbed (dashed) air mass in a long rigid pipe.

The pipe has an inner cross-sectional area $S = \pi a^2$. The undisturbed air is that contained in the cylindrical volume between the planes defined by z and $z + \Delta z$. The mass, m, of this air is the static air density multiplied by the volume of the cylinder between the solid lines.

$$m = \rho_0 S\Delta z \qquad (16\text{-}11)$$

Imagine now that a closely fitting piston is inserted into the pipe on the left and quickly displaces the air particles that were originally on the plane at z to the new dashed planar position $z + \xi$ such that all of the air particles that were originally at z are now located at $z + \xi$. In other words, the air particles originally located at the spatial coordinate z have undergone an amount of displacement equal to ξ. Note also that the particle displacement does not depend on the spatial coordinates r or θ. All air particles having a particular value of the z coordinate are displaced the same amount such that the particle displacement depends only on z and on time, t. Now if air were incompressible, all of the particles originally on the plane at $z + \Delta z$ would be displaced to a new plane at $z + \Delta z + \xi$. Air is compressible however, so we must allow for the particle displacement to undergo a change over the space interval of Δz so that the right extremity of our disturbed mass of air is located at $z + \Delta z + \xi + \Delta\xi$. Our original mass of air is now contained in the cylinder defined by the two dashed planes. The air has been compressed as a result of the piston motion. As a consequence, the volume of the disturbed cylinder is slightly less than that of the undisturbed one and this simply requires that $\Delta\xi$ be a negative number. The mass of air was conserved in the process so the density of air in the disturbed cylinder has increased. The volume of the undisturbed cylinder is $S\Delta z$ while that of the disturbed cylinder is $S(\Delta z + \Delta\xi)$ so we may write

$$\begin{aligned} \rho_0 S\Delta z &= \rho S(\Delta z + \Delta\xi) \\ &= \rho S\Delta z\left(1 + \frac{\Delta\xi}{\Delta z}\right) \end{aligned} \qquad (16\text{-}12)$$

This equation readily simplifies to

$$\rho_0 = \rho\left(1 + \frac{\Delta\xi}{\Delta z}\right) \qquad (16\text{-}13)$$

In words this last equation says that the undisturbed density of air is equal to the disturbed density multiplied by one plus the average slope of the particle displacement function over the interval Δz. This slope is negative however as the particle displacement decreases as z increases and thus the number in the parenthesis is less than one. The average slope is not good enough. We need to make our calculation independent of our choice of the size of Δz. At this point I recognize that many readers have not had an opportunity to study calculus much less partial differential equations. Both of these are required in order to do a rigorous derivation of the wave equation. In much of the following then, I will substitute word descriptions for what is going on rather than adhering to pure mathematical formalism. In the density equation above we take successively smaller and smaller sizes for Δz or in other words let Δz approach zero all the while studying the ratio $\Delta\xi/\Delta z$ and look to see what limiting value is approached by the quotient of $\Delta\xi/\Delta z$. This limit is called the partial derivative of the particle displacement with respect to the z coordinate and the density relation is then written as

$$\rho_0 = \rho\left(1 + \frac{\partial\xi}{\partial z}\right) \qquad (16\text{-}14)$$

The reason for doing this is to find the value of the disturbed air density in the immediate vicinity of the point z and at time t. Among other things, we want to learn how the air density behaves under disturbed conditions as a function of position and time. This last equation tells us how to calculate the density behavior once we know how the particle displacement behaves. Recall that the condensation is given by

$$s = \frac{\rho - \rho_0}{\rho_0} \qquad (16\text{-}15)$$

This can be solved for the disturbed density to yield $\rho = \rho_0(1 + s)$. This is now substituted in the density relation to yield

$$\rho_0 = \rho_0(1 + s)\left(1 + \frac{\partial\xi}{\partial z}\right) \qquad (16\text{-}16)$$

$$1 = (1 + s)\left(1 + \frac{\partial\xi}{\partial z}\right) \qquad (16\text{-}17)$$

$$1 = 1 + \frac{\partial\xi}{\partial z} + s + s\frac{\partial\xi}{\partial z} \qquad (16\text{-}18)$$

In order for the remainder of our development to be as simple as possible we must restrict the size of the acoustical disturbance to that for which the air behaves as a linear medium. Even with this restriction the equations that we develop will accommodate sound pressure levels up to 120 dB with little error. With this restriction, we can observe that in the last equation written above both s and the partial derivative of the displacement with respect to z are small quantities individually and that the product of the two of them is very small indeed. Hence neglecting the product term introduces negligible error. This final equation can then be rewritten as

$$s = -\frac{\partial \xi}{\partial z} \quad (16\text{-}19)$$

Previously we learned that for small disturbances the acoustic pressure is given by

$$p = c^2(\rho - \rho_0) \quad (16\text{-}20)$$

and in terms of the condensation this may be written as

$$p = \rho_0 c^2 s \quad (16\text{-}21)$$

Alternatively, we may express the acoustic pressure as

$$p = -\rho_0 c^2 \frac{\partial \xi}{\partial z} \quad (16\text{-}22)$$

One other observation is appropriate at this point. The particle displacement, ξ, is in general a function of both z and t. This means that $\xi = \xi(z,t)$. In fact, one of our objectives is to find the exact nature of this function for a given type of acoustical excitation. Once we determine the nature of this function we can determine the value of the particle displacement for any value of the spatial coordinate z and time coordinate t. Additionally we will also be able to determine the particle velocity, u, as the particle velocity at any particular value of the z coordinate and time t is the rate at which the particle displacement is changing with time at the fixed value of z. The particle velocity is given by the partial derivative of the particle displacement with respect to time and is written as

$$u = \frac{\partial \xi}{\partial t} \quad (16\text{-}23)$$

Similarly, the local particle acceleration or rate of change of velocity is calculated from

$$\frac{\partial u}{\partial t} = \frac{\partial^2 \xi}{\partial t^2} \quad (16\text{-}24)$$

Thus far we have required only a few definitions, the law of conservation of mass, and knowledge of the behavior of air while undergoing small adiabatic compression or expansion. Now with the help of Sir Isaac Newton's second law of motion we will be able to finally arrive at the plane wave equation. First we must list the forces that could possibly affect the air particle motion in the tube. The principal force is that exerted by the piston as it first begins to compress the air at the left face of our undisturbed cylinder of air as depicted in Fig. 16-7 The pressure exerted by the piston must exceed static atmospheric pressure in order to produce compression so we write this as $(P_0 + p)S$ where p is the acoustic pressure at z. Similarly, the pressure at the right face is the static pressure plus the acoustic pressure at the right face which we must allow to be different from that at the left face. We write the force at the right face then as $(P_0 + p + \Delta p)S$. In principle, the force of gravity would tend to make the static pressure at the bottom of the tube minutely greater than that at the top. This effect is insignificant for tubes of ordinary diameters. Finally, we should mention the possibility of viscous effects. Viscous frictional forces occur principally at the tube walls and manifest themselves as a small attenuation in very long tubes. We will neglect such effects for reasons of simplicity.

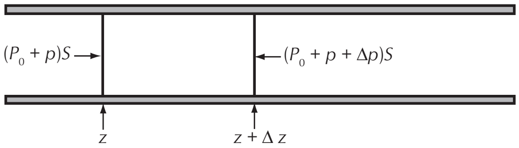


Figure 16-7. Forces acting on undisturbed element at the onset of compression by the piston. S is the cross-sectional area of the pipe.

From Fig. 16-7, the net force in the positive z direction is $-\Delta pS$. According to Newton's second law the net force acting on the mass of air in the element must be equated to the mass multiplied by the acceleration or

$$-\frac{\Delta p}{\Delta z}\Delta z S = \rho_0 \Delta z S \frac{\partial^2 \xi}{\partial t^2} \quad (16\text{-}25)$$

Upon canceling common factors and taking the limit as we have done in a previous case this equation becomes

$$-\frac{\partial p}{\partial z} = \rho_0 \frac{\partial^2 \xi}{\partial t^2} \quad (16\text{-}26)$$

In words this result says that the negative of the space rate of change of the acoustic pressure at a given point and time is the undisturbed density of air multiplied by the particle acceleration at the same space point and time t.

Two more steps and we will be at the punch line. One of our previous results while studying the particle displacement was

$$p = -\rho_0 c^2 \frac{\partial \xi}{\partial z}$$

From this relation we need to calculate the space rate of change or the slope of the acoustic pressure. This is done by calculating the partial derivative with respect to the z coordinate on both sides of the equation. The result is

$$\frac{\partial p}{\partial z} = -\rho_0 c^2 \frac{\partial^2 \xi}{\partial z^2} \tag{16-27}$$

This is now substituted into the equation derived employing Newton's second law to produce

$$\frac{\partial^2 \xi}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2 \xi}{\partial t^2} \tag{16-28}$$

This second order partial differential equation is the governing equation for plane waves that depend on only one space coordinate and time as the independent variables. The dependent variable in this instance is the air particle displacement. Instead of concentrating on the particle displacement as the dependent variable, we could have just as well done a parallel development while centering our attention on the acoustic pressure to obtain

$$\frac{\partial^2 p}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} \tag{16-29}$$

In other words, the acoustic pressure and the particle displacement are governed by the same partial differential equation. In order for some mathematical function to be a solution to a physical circumstance involving the plane wave equation it must accomplish three things. Firstly, when substituted into the wave equation it must yield an identity. Secondly, it must satisfy the conditions that exist at $t = 0$. Finally, it must satisfy the conditions that exist at the coordinate boundaries for all values of $t \geq 0$. There are many functions that satisfy the first condition. In fact, there are an infinite number of such functions. All of these functions, however, have one feature in common and that is whenever the space and time independent variables appear in one of the functions, this appearance must be of the form $(ct \pm z)$. The two other requirements play the role of sorting through this infinite set to find the one and only solution that fits the problem at hand. We are assured that there is only one genuine solution to the wave equation that satisfies the three stated requirements because of the existence of a uniqueness theorem governing solutions to the wave equation. In order to make this really meaningful, we must seek a solution to this equation for a realizable physical circumstance. First, let's illustrate the significance of $(ct \pm z)$. Suppose we have a very long plane wave tube with the origin of coordinates at the mid-point of the tube. Further suppose at $t = 0$ that some disturbance produces an acoustic pressure matching only the solid curve in Fig. 16-8.

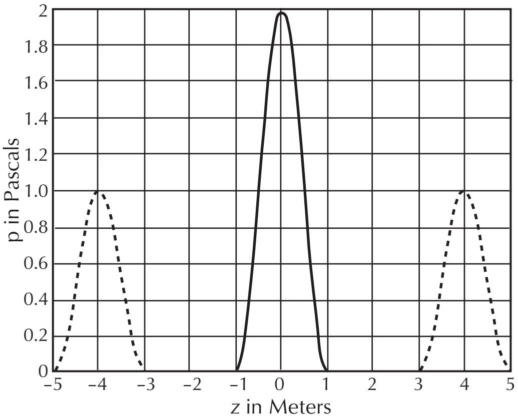


Figure 16-8. Plane wave tube with an initial disturbance (solid) at its center.

Refer now to Fig. 16-8 and imagine that only the solid curve is present. This would represent the first frame of a movie describing the acoustic pressure versus time and position in space. The second frame would show the initial disturbance beginning to split into two equal parts with one part displaced slightly to the left and the other displaced an equal amount to the right. Many frames later, the solid curve would no longer be present and the two dashed curves would represent a snapshot at the instant when $ct = 4m$. In other words the initial static disturbance has evolved into two traveling disturbances moving in opposite directions along the z-axis. The functions involved would be

$$p_+ = p(ct - z) \tag{16-30}$$

and

$$p_- = p(ct + z) \tag{16-31}$$

Let's concentrate on just the p_+ term. If we are to always observe the same peak pressure for this term, what must we do as an observer? Remember that we have no control over time. It increases uniformly whether we want it to or not. As ct increases uniformly then we must increase our location on the z-axis at the same rate such that $ct - z$ maintains a value of, in this case, zero. The rate at which ct increases is the speed of sound, c. Therefore an observer must race in the direction of increasing z with a speed equal to c in order to keep up with the pressure pulse on the right. Similarly, an observer must race in the direction of decreasing z with a speed c in order to keep up with the pressure pulse on the left.

Thus we have the plane wave equation in hand and have learned the properties that must be exhibited by a mathematical function if it were to be a solution to the plane wave equation for a given set of physical circumstances. It is now time to apply what we have learned towards the study of the wave motion in a plane wave tube that is excited by an oscillating piston at one end as depicted in Fig. 16-9.

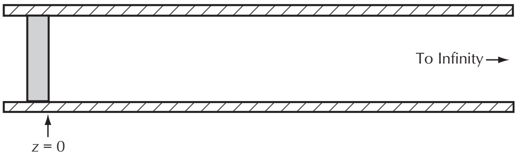


Figure 16-9. Plane wave tube that is fitted with an oscillating piston.

For simplicity, we will consider the circumstance where the piston, depicted in gray, has been forced to undergo oscillatory motion for some time by a mechanism not shown in the figure and continues to do so as we study the problem. We will start measuring time from the instant when the right face of the piston is just passing $z = 0$ and is moving to the right such that the piston's displacement from $z = 0$ is described by

$$\xi = \xi_m \sin(\omega t) \tag{16-32}$$

where,

ξ_m is the amplitude of the piston displacement,

$\omega = 2\pi f$ with f being the frequency of oscillation in Hz,

t is the time.

Now we want to find a solution to the wave equation for air particle displacement in the tube that satisfies these conditions. The tube is infinitely long so there can be no reflections from the receiving end. As a consequence, we need only a solution that describes a wave traveling to the right. The air in contact with the piston undergoes the same motion as does the piston itself so we propose as a solution an expression that duplicates the piston motion when we let $z = 0$

$$\xi = \xi_m \sin(\omega t - kz) \tag{16-33}$$

We also have learned that in order for a solution to legitimately describe a plane wave propagating in the direction of increasing z that the space and time variables must appear in the form $(ct - z)$. We can easily show that our proposed solution satisfies this requirement as follows. The quantity k is called the propagation constant and is defined as $k = 2\pi/\lambda$ where λ is the wavelength. Now as ω is $2\pi f$ and ω/k is $\lambda f = c$, then if we factor k out of our parenthesis in our proposed solution, the solution will take the form

$$\xi = \xi_m \sin[k(ct - z)] \tag{16-34}$$

Since the two expressions for the particle displacement are equivalent we may use either form to suit our convenience. Next, it is necessary to show that our proposed solution when substituted into the wave equation produces an identity. In accomplishing this it is necessary to take partial derivatives of our proposed solution first with respect to z and then with respect to t. The first partial derivative with respect to z finds the slope of the solution when z is allowed to change while t is held at a constant value. Similarly, the second partial derivative with respect to z finds the slope of the slope curve while z is allowed to change with t being held constant. The process is then repeated except now t is allowed to change while z is held at a constant value. The results are found to be

$$\frac{\partial^2\xi}{\partial z^2} = -k^2\xi_m \sin(\omega t - kz) \tag{16-35}$$

$$\frac{\partial^2\xi}{\partial t^2} = -\omega^2\xi_m \sin(\omega t - kz) \tag{16-36}$$

The wave equation tells us to divide the second partial derivative with respect to t by c^2 and equate the result to the second partial derivative with respect z. If an identity results from this action then our proposed solution does indeed satisfy the wave equation. Upon dividing the second equation immediately above by c^2 and equating it to the first equation immediately above we obtain

$$-\frac{\omega^2}{c^2}\xi_m \sin(\omega t - kz) = -k^2\xi_m \sin(\omega t - kz) \tag{16-37}$$

This is indeed an identity because $k = 2\pi/\lambda = 2\pi f/(c) = \omega/c$. Finally, when we let $z = 0$, our air particle displacement agrees with the piston motion at all times t including $t = 0$. Therefore our proposed solution satisfies all of the requirements necessary to be the one and only solution to the problem.

What about the particle velocity and the acoustic pressure? We obtain the particle velocity from the partial derivative of the particle displacement with respect to t.

$$u = \frac{\partial\xi}{\partial t} = \omega\xi_m \cos(\omega t - kz) \tag{16-38}$$

The acoustic pressure is obtained from $(-(\rho_0))c^2((\partial\xi)/(\partial z))$. We learned this in the beginning of this chapter.

$$p = \rho_0 c^2 k \xi_m \cos(\omega t - kz) \tag{16-39}$$

It is important to note that if we divide the acoustic pressure expression by that of the particle velocity we obtain a quantity called the specific acoustic impedance of air for plane waves namely,

$$\frac{p}{u} = Z_s = \frac{\rho_0 c^2 k}{\omega} = \rho_0 c \qquad (16\text{-}40)$$

Here the capital letter Z represents impedance rather than the spatial coordinate and the subscript s stands for specific. The specific acoustic impedance of air for plane waves is a real number denoting the fact that the acoustic pressure and particle velocity are in phase. The dimensions of Z_s are kg•m^{-2}•s^{-1}. This combination is called a Rayl in honor of Lord Rayleigh who was a pioneer in the study of sound and acoustics.

Now we will put the theory into practice with a realistic numerical example. Let the frequency of oscillation of the piston be 1000 Hz and let its displacement amplitude be 10^{-6} m. Let the static air pressure be the sea level value but let the temperature be a comfortable 70° F. This corresponds to 21.11° C or 294.26 K. The static air density is inversely proportional to the absolute temperature so then $\rho_0 = 1.293(273.15/294.26 = 1.20$ kg•m^{-3}. The speed of sound is directly proportional to the square root of the absolute temperature so $c = 331.46(294.26/273.15)^{0.5} = 344$ m•s^{-1}. The air particle displacement amplitude matches that of the piston so $\xi_m = 10^{-6}$ m. The angular frequency $\omega = 2\pi f = 6283$ radians/s. The propagation constant $k = \omega/c = 18.265$ m^{-1}. The velocity amplitude is $u_m = ck\xi_m = \omega\xi_m = 6.283(10^{-3})$ m•s^{-1}. The acoustic pressure amplitude is $p_m = \rho_0 c u_m = 2.5937$ Pa. The rms pressure for sinusoidal time dependence is the amplitude multiplied by 0.7071 and is 1.834 Pa. This corresponds to a *SPL* of 99.25 dB. The wavelength $\lambda = c/f = 0.344$ m. Our solutions for the acoustic variables expressed as functions of both position and time are then

$$\xi = 10^{-6} m \bullet \sin\left(\frac{6,283}{\text{sec}}t - \frac{18.265}{m}z\right) \qquad (16\text{-}41)$$

$$u = 6.283(10^{-3}) m \bullet \text{sec}^{-1} \bullet \cos\left(\frac{6,283}{\text{sec}}t - \frac{18.265}{m}z\right) \qquad (16\text{-}42)$$

$$p = 2.5937 Pa \bullet \cos\left(\frac{6,283}{\text{sec}}t - \frac{18.265}{m}z\right) \qquad (16\text{-}43)$$

Given that the piston has been oscillating for some time, Fig. 16-10 depicts the acoustic pressure wave propagation along a one-wavelength interval of the z-axis versus elapsed time commencing from the instant when the piston is located at $z = 0$ and is moving in the positive z direction.

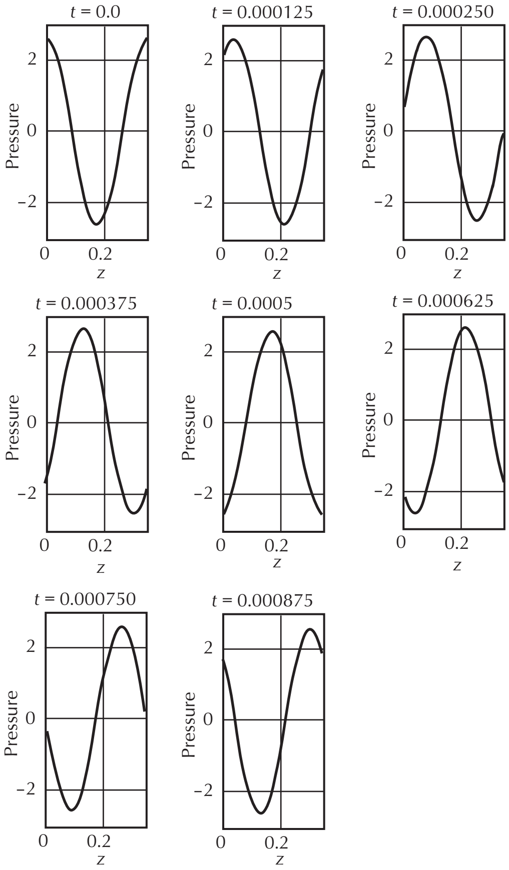


Figure 16-10. A depiction of successive shifts of the pressure waveform along the z-axis as time increases.

Now for a pop quiz! If we were to construct a ninth entry to Fig. 16-10 corresponding to $t = 0.001$ sec, how would it look? Hint: 0.001 sec corresponds to the period of the motion of the piston and is equal to the time required for the pressure wave to travel a distance of one wavelength along the z-axis. This being the case, the ninth entry would look exactly like the first. Furthermore, a slight modification of Fig. 16-10 would allow it to describe the particle velocity as well. This modification would involve only a change of scale and label for the vertical axes as the particle velocity is in phase with the acoustic pressure for a plane wave in air.

Now that we have established the behavior of the sound wave in the tube it is appropriate to consider what the piston's motion must accomplish to bring about this behavior. The piston of course is displacing the air adjacent to its right hand face. The piston must exert a force on the air in order to displace it and this requires that the piston perform work on the air. The force, *F*, exerted by the piston at any instant is the acoustic pressure at $z = 0$ multiplied by the cross-sectional area of the tube namely, *S*.

$$F = Sp_m \cos(\omega t) \tag{16-44}$$

The rate at which the piston is performing work on the air is the instantaneous power or *P* and is obtained by multiplying the applied force by the rate of displacement at $z = 0$. The rate of displacement at $z = 0$ is just the particle velocity at the origin so

$$\begin{aligned} P &= Sp_m \cos(\omega t) \bullet u_m \cos(\omega t) \\ &= Sp_m u_m \cos^2(\omega t) \end{aligned} \tag{16-45}$$

Fig. 16-11 is a plot of this result for one period of the piston motion using the values from our numerical example when applied to a plane wave tube having an inner diameter of 1 in or 0.0254 m.

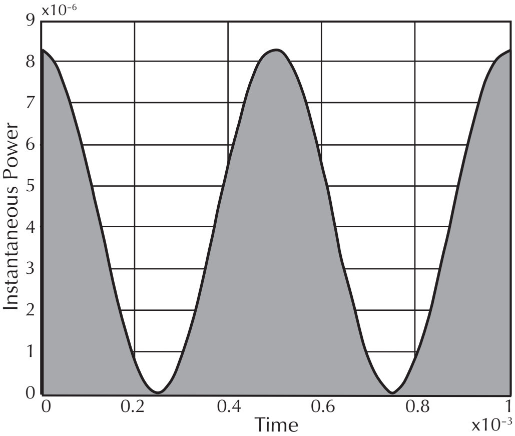


Figure 16-11. Instantaneous power delivered by the piston to the air in the plane wave tube. The grey area is the acoustic energy delivered to the sound wave in one period of the piston's motion.

Fig. 16-11 displays two items of interest: the plot of the instantaneous power versus time and the area beneath the power curve that is shaded gray. Since the average value of the $\cos^2$ over one period is 1/2 the area under the curve is $1/2 P_m$•0.001 sec. For our example this would be $4.1288(10^{-9})$ J. This area accounts for the total acoustic energy delivered to the sound wave during one period of the piston's motion. One can reasonably inquire as to where this energy resides in the sound wave. The acoustical energy associated with a plane wave appears in two forms. First there is acoustic kinetic energy associated with the motion of the air particles themselves and then there is acoustic potential energy associated with the existence of acoustic pressure. We encountered the concept of acoustic potential energy earlier in this chapter. This acoustic energy is not localized at a point but rather is distributed throughout the volume occupied by the wave with an energy density that varies as a function of position and time. If we let *e* represent the total acoustic energy density while e_k and e_p represent the kinetic and potential energy densities, respectively, then

$$\begin{aligned} e &= e_k + e_p \\ &= \left(\frac{1}{2}\rho_0 u^2 + \frac{1}{2}\frac{p^2}{\rho_0 c^2}\right) \\ &= \frac{p^2}{\rho_0 c^2} \end{aligned} \tag{16-46}$$

The last step in the above equation is justified because for a plane wave in air the particle velocity and acoustic pressure are related through $u = p/(\rho_0 c)$ thus making the kinetic and potential energy densities equal with each being one-half of the total energy density. The total acoustic energy density is also a function of position and time. Using the data from our numerical example the total acoustical energy density expression becomes

$$e = 4.7374(10^{-5})\mathrm{J} \bullet \mathrm{m}^{-3} \bullet \cos^2\left(\frac{6,283}{\mathrm{sec}}t - \frac{18.265}{\mathrm{m}}z\right) \tag{16-47}$$

Fig. 16-12 is a plot of this energy density for a one-wavelength interval along the z-axis at the instant when $t = 0.001$ sec. This corresponds to an elapsed time of one period of the piston motion.

Now we are in a position to calculate the total acoustic energy contained in the plane wave tube for a one-wavelength interval along the z-axis. If we draw a horizontal line across the peaks of the curve, we will now have a rectangle whose area numerically is $0.344 \bullet 4.7374(10^{-5})$. By visual inspection, however, the actual area beneath the curve indicated in gray is only 1/2 of this value so the average height of the curve is 1/2 of its peak value meaning that the average value of the energy density in this interval is $2.3687(10^{-5})\,\mathrm{J} \bullet \mathrm{m}^{-3}$. (This is just an illustration of the fact that the average value of $\cos^2$ over one period is 1/2.) Here is the punch line. The total energy in the wave for this one-wavelength interval is the average energy density in the wave multiplied by the volume occupied by the wave. This is $S\lambda\langle e\rangle$ where $\langle e\rangle$ is the average value of the acoustic energy density. Calling this energy *W*, we have

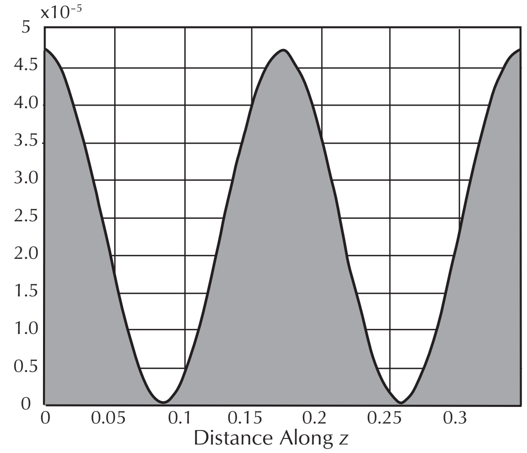


Figure 16-12. Acoustic energy density at $t = 0.001$ sec for a one-wavelength interval along the z-axis.

$$\begin{aligned} W &= S\lambda\langle e\rangle \\ &= \pi\left(\frac{0.0254\ \text{m}}{2}\right)^2 0.344\ \text{m}(2.3687)10^{-5} \\ &= 4.1288(10^{-9})\ \text{J} \end{aligned}$$

This is just the amount of energy supplied by the piston in the previous 0.001 sec!

Now if an observer is positioned at any fixed value of the z-coordinate in the plane wave tube, this same amount of acoustical energy will pass the observation point while being transported in the positive z direction in a time of 0.001 sec so the average acoustical power over this interval of time is W/T where W is the acoustical energy and T is the period of the sinusoidal piston motion. $\langle P\rangle$ denotes this average power and in this instance has the value 4.1288 (10^{-6}) Watt. The average acoustical intensity denoted by $\langle I\rangle$ is a vector quantity defined to be the average directed power flow per unit area. As it is a vector quantity $\langle I\rangle$ has both a magnitude and direction. For a fixed location the magnitude of $\langle I\rangle$ amounts to the total acoustical energy flow averaged over the time of the flow per unit of area through which the energy passes. In this instance the magnitude of $\langle I\rangle = \langle P\rangle/S$ where S is the cross-sectional area of the tube. In the present case the magnitude of $\langle I\rangle$ is $8.1482(10^{-3})$ Watt•m^{-2} and the direction of $\langle I\rangle$ is the same as that of the wave propagation, namely the positive z-direction. The instantaneous intensity $I(z,t)$ is a related physical quantity that is also a vector quantity. $I(z,t)$ is a function of both position and time and is a measure of the instantaneous power flow per unit area at a particular location z and time t. It is calculated from the product of the acoustic pressure with the particle velocity at the z-coordinate and time coordinate of interest so $I(z,t) = p(z,t)\bullet u(z,t)$. For a plane wave propagating in the direction of increasing z, $u(z,t) = p(z,t)/\rho_0 c$. For the case at hand, the direction of $I(z,t)$ is always that of the positive z-axis. It is true that the particle velocity alternates between the positive and negative z direction, but the acoustic pressure is in phase with the particle velocity so that when the particle velocity is instantaneously in the negative z direction the acoustic pressure is negative and the overall product remains positive. Now $\langle I\rangle$ at some fixed point z can be calculated from the time average of $I(z,t)$ at the same value of z and hence $\langle I\rangle = \langle[p(z,t)]^2\rangle/(\rho_0 c)$. The quantity $\langle[p(z,t)]^2\rangle$ by definition is just the mean value of the square of the acoustic pressure so it is possible to write this equation as

$$\langle I\rangle = \frac{p_{rms}^2}{\rho_0 c} \tag{16-48}$$

This is a very useful equation. Even though it was derived by considering only a plane wave it is equally valid for a spherical wave.

It is probably safe to say that the most often made measurement in acoustics is that of sound pressure level. Levels are logarithmic comparisons of a power or power-like quantity with regard to some standard reference value for the quantity in question. When the quantity is acoustic pressure, the power-like quantity is p_{rms}^2 and the reference is $[20(10^{-6})]^2$ Pa^2. Strictly speaking, the sound pressure level in decibels would be written with the form

$$\begin{aligned} SPL &= 10\ \text{dB}\ \log_{10}\left(\frac{p_{rms}}{20(10^{-6})}\right)^2 \\ &= 20\ \text{dB}\ \log_{10}\left(\frac{p_{rms}}{20(10^{-6})}\right) \end{aligned} \tag{16-49}$$

We have previously learned that the root mean square acoustic pressure for the sound wave in this plane wave tube has a value of 1.834 Pa. When this value is substituted into the equation for *SPL*, the calculated sound pressure level becomes 99.25 dB. Additionally, when p_{rms} is substituted along with $\rho_0 c$ having a value matching the ambient conditions, namely 412.8 Rayls, the magnitude of $\langle I\rangle$ is found to be $8.1482(10^{-3})$ Watt•m^{-2}. Now $\langle I\rangle$ is a true power-like quantity and the reference value for average acoustic intensity is 10^{-12} Watt•m^{-2}. The intensity level or *IL* is then calculated from

$$\begin{aligned} IL &= 10\ \text{dB}\ \log_{10}\left(\frac{\langle I\rangle}{10^{-12}}\right) \\ &= 99.1\ \text{dB} \end{aligned}$$

There are two reasons for the small discrepancy in a given physical situation between the numerical values for *SPL* and *IL*. Firstly, the reference values, although close, are not exactly equivalent and the specific acoustic impedance of air for plane waves varies dependent upon the ambient total atmospheric pressure and the absolute temperature. For the range of ambient conditions usually encountered in practice they will agree within a fraction of a decibel as was the case in our example. For all practical purposes, then, the *SPL* and the *IL* can be taken as one and the same.

Plane wave tubes are often used for measuring the properties of transducers in general and high frequency compression drivers in particular. We have seen that a uniformly constructed tube of infinite length with excitation at one end allows the existence of a single traveling plane wave in just one direction. Such a device is obviously not physically possible. We need to visualize a device of finite length that maintains uniform geometry and does not produce reflections because of its finite length. Two such possibilities are suggested in Figs. 16-13 and 16-14.

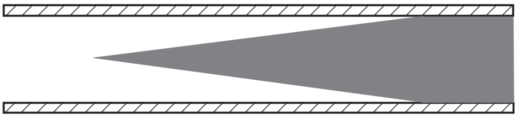


Figure 16-13. Possible plane wave tube structure.

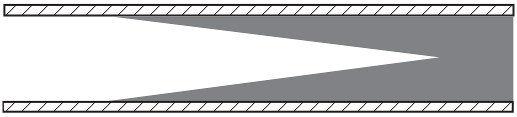


Figure 16-14. Alternative plane wave tube structure.

Figs. 16-13 and 16-14 are cross-sectional drawings of viable plane wave tube structures. In each instance the inner tube diameter is greatly exaggerated as compared with the actual tube length. Typical inner diameters for use with compression drivers must exactly match the exit apertures of the drivers of interest. The inner diameters would then be of the order of an inch or so but the tube must be ten or more feet in length. In both instances the structure is a figure of revolution about the central axis so that cylindrical symmetry is maintained over the entire length of the structure. As a result there are no abrupt changes in geometry. The interior shaded regions in both instances are occupied by a uniform open cell acoustical foam that has a specific acoustical impedance that matches as closely as possible that of air in the range of 410 to 415 Rayls. The tubes must be mounted vertically because acoustical foam is flexible and a horizontal mount would lead to sag of the foam that would destroy the cylindrical symmetry. This would be particularly true for the conical foam structure of Fig. 16-13. The hollow tube employed in the structure in both instances must have smooth interior walls of sufficient thickness to be rigid.

The rationale behind the proposed structure is quite straightforward. Having an exact match between driver exit diameter and the tube's inner diameter ensures two things. In the first instance the driver's compression ratio will not be influenced by its attachment to the tube and secondly, there will not be an abrupt change in geometry that would cause a reflection back into the driver. In the air-filled space near the driver the plane wave propagates in the normal fashion and as the wave progresses down the tube it encounters no abrupt geometry changes and is always in media having a common value of specific acoustical impedance. The portion of the wave in the free air is undergoing an adiabatic process and loses no acoustical energy. The portion of the wave in the acoustical foam on the other hand is in a medium of much higher thermal conductivity and is undergoing a predominantly isothermal process where acoustical energy is being dissipated as heat. Finally the greatly attenuated wave enters a region entirely filled with foam for a sufficient length that the remaining acoustical energy for all practical purposes is completely absorbed by the time the end of the tube is reached. In summary, there are no reflections back towards the source. The downside associated with these structures is the necessity of the vertical orientation, where ceiling heights are restricted, and the expense involved in accurately shaping the acoustical foam.

As an alternative, Fig. 16-15 depicts what might be called a poor man's plane wave tube. In viewing Fig. 16-15 the reader should be aware that the dimensional scales associated with the driver mounting flange, the microphone port, and the tube diameter are greatly magnified relative to the tube length in order to show construction details. The microphone port should be as close to the driver as possible and should make a tight fit with the body of a pressure sensitive microphone assumed to be cylindrical in shape. The microphone should be no more than 0.5 in in diameter with a 0.25 in diameter microphone being preferred. In either case the microphone capsule's protective grid should be removed and the microphone should be positioned such that its diaphragm's surface is just tangent with the inner wall of the plane wave tube. Half of the length of the tube is to be filled with graduated stuffing of ordinary fiberglass building insulation. The shaded interior portion of the drawing in Fig. 16-15 indicates this. The graduated stuffing is prepared in the following way. The final one-foot length near the end of the tube should be compacted firmly and then the amount of compaction should be gradually reduced until the midpoint of the tube is reached. Thick-walled PVC piping of the appropriate inner diameter may be employed for the plane wave tube and a smaller pipe with a plugged end with incremental length markings can be employed as a stuffing plunger while working with small tuffs of the fiberglass. The principal tube may be mounted horizontally if it is provided with adequate supports to keep the tube level. Several squares of ¾ in plywood with an appropriate center hole sized to fit the outside diameter (o.d.) of the plane wave tube are adequate. The graded absorber is admittedly not perfect. If care is taken in making it, however, any reflections will be close to 40 dB down. If the distance between the microphone port and the face of the graded absorber is at least 5 ft then the first possible reflection will return to the measuring port after a time of 10 ft/1128 ft/s or approximately 0.009 s. Furthermore, if one employs a TEF, Sysid, or similar measurement program the contribution from any reflections can be screened out with only a moderate loss in the frequency resolution of the measurement. If sufficient horizontal space is available, a tube of 20 ft of overall length with 10 ft of graduated stuffing will allow accurate measurement to as low as 50 Hz.

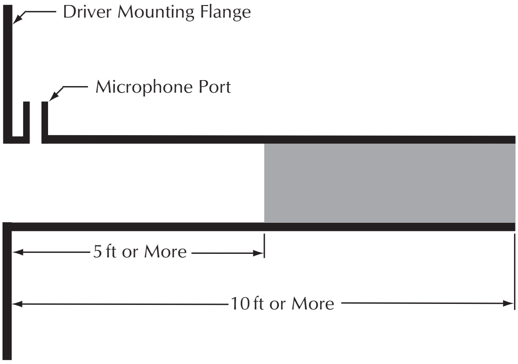


Figure 16-15. Poor man's plane wave tube for horizontal mounting.

There is one caution related to high frequency operation. A cylindrical wave guide can support modes of wave propagation other than that of plane waves if the guide is suitably excited above a frequency as given by the following equation, $f = 1.84c/(\pi d)$ where d is the inner diameter of the wave-guide.

This frequency is approximately 8 kHz for a one-inch diameter guide and of course 4 kHz for a two-inch diameter guide. This matter as well as wave-guide behavior for various impedance miss-matches will be the topic of a later discussion.

16.3 Non-Planar Wave Motion in a Tube

When a plane wave tube is excited at its origin by a tightly fitting, oscillating piston as illustrated in Fig. 16-16, the resulting wave motion is that of a plane wave propagating in the direction of increasing z. In this motion, the acoustic pressure and the particle velocity are uniform over the cross-section of the tube and the particle velocity oscillates only in the z-direction. The phase velocity of this plane wave motion is independent of the frequency of excitation. This is not the case for an arbitrary type of excitation nor is it necessarily true when a compression driver excites a plane wave tube as the emerging wave front from such a device may have some curvature. In such instances, one must consider a more general solution to the wave equation consistent with the geometry of the plane wave tube.

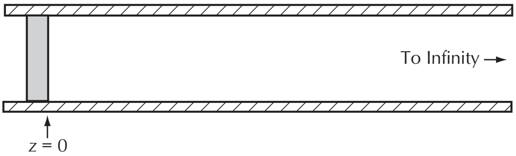


Figure 16-16. Plane wave tube excited by an oscillating piston at the origin.

In terms of the cylindrical coordinates, (r, θ, and z), that are the simplest ones to employ for the geometry at hand the general solution to the wave equation for acoustic pressure is expressed as a product of three different functions. The first of these functions describes how the acoustic pressure depends on the radial distance from the central or z-axis of the tube. The second function describes how the acoustic pressure varies with the polar angle measured about the central axis. The third function describes how the acoustic pressure varies with regard to both position along the z-axis and with time.

The radial behavior is described by a Bessel function of the first kind of which there are many choices depending upon exactly what mode of wave motion is involved. These Bessel functions are ordered by a subscript m. Bessel functions with orders 0 through 3 are depicted in Fig. 16-17.

As can be seen from Fig. 16-17 the Bessel functions of the first kind appear almost as damped sine or cosine functions of the variable x although they are not, as the zero crossings are not periodic. One needs to refer to math tables or computer based math programs to obtain detailed behaviors. The variable x employed in Fig. 16-17 does not refer to the space variable x but rather to the combination $k_{mn}r$ where r is the radial distance from the z-axis and k_{mn} is the

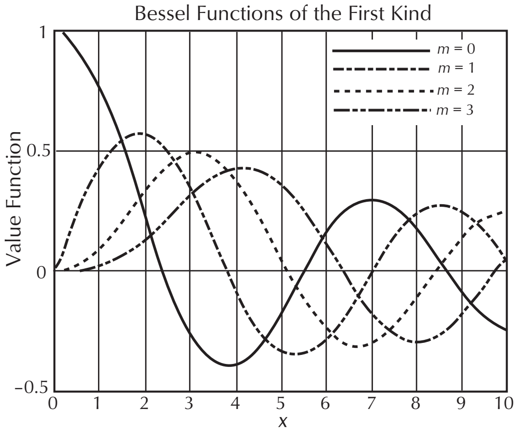


Figure 16-17. Bessel functions $J_m(x)$ for orders $m = 0$ through $m = 3$.

radial wave motion propagation constant. The radial propagation constant k_{mn} requires some extended discussion. First off it has two integral indices m and n. The index m refers to the order of the Bessel function while the index n refers to the order of the position of the variable x in Fig. 16-17 where the particular Bessel function at hand has zero slope. This zero slope is important because an acceptable solution can only be one for which the radial component of the particle velocity must vanish at the rigid wall of the waveguide and this radial component of particle velocity is proportional to the derivative of the acoustic pressure with respect to the variable r. When $r = a$, the derivative of the pressure with respect to r, that is the slope, must be zero. For example, let $m = 1$. The first value of x beyond the origin at which the slope of this curve is zero is at the point $x = 1.841$. This requires then that k_{11} must be $1.841/a$ in order for $k_{11}r = 1.841$ when r becomes equal to a. Similarly, when $m = 2$, the first occurrence of zero slope is for $x = 3.054$ requiring k_{21} be $3.054/a$. The significance of this can be learned from the relationship between the radial propagation constant k_{mn} and the propagation constant along the z-axis that is k_z. This relationship is $k_z = [(\omega/c)^2 - (k_{mn})^2]^{1/2}$. In order to have a propagating mode along the z-axis, k_z must be a real number. This will be true only for those operating frequencies where $(\omega/c)^2 > (k_{mn})^2$ or when $\omega > k_{mn}c$. Consider the mode where $m = 1$ and $n = 1$. The frequency below which this mode cannot propagate, that is the cutoff frequency, is given by

$$f_{11} = (1.841c)/(2\pi a)$$

and

$$f_{21} = (3.054c)/(2\pi a)$$

Table 16-3 lists all of the modal cutoff frequencies in the 20 kHz audio band for a one-inch diameter (0.0254 m) plane wave tube. The cutoff frequencies for a two-inch diameter tube are one-half those for a one-inch diameter tube.

Table 16-3. Cutoff Frequencies in the Audio Band for a One-inch Diameter Plane Wave Tube

m	n	k_{mn}	f_{mn} in Hz
1	1	1.841 / a	7936
2	1	3.054 / a	13,166
0	2	3.832 / a	16,520
3	1	4.20 / a	18,106

The modes listed in Table 16-3 are dispersive modes. They propagate at operating frequencies above their respective cutoff frequencies, but do so with a frequency dependent phase velocity. This means that different frequency components of a wideband signal above cutoff propagate with different speeds and wave shapes are not preserved. The phase velocity measured along the z-axis is given by $\omega/[(\omega/c)^2 - k_{mn}{}^2]^{1/2}$. Below its respective cutoff frequency, each mode becomes evanescent. This means that the mode does not propagate or transport energy along the z-axis but rather its pressure contribution attenuates exponentially with distance from the origin.

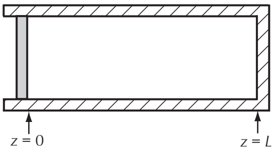
A reasonable question to ask at this point is, “What does the solution look like with all of these added complications?” The answer is a sum over all indices that can contribute for a given operating frequency or range of operating frequencies of pressure terms of the following structure

$$p_{mn}(r, \theta, z, t) = A_{mn}J_m(k_{mn}r)\cos(m\theta)\cos(\omega t - k_z z) \tag{16-50}$$

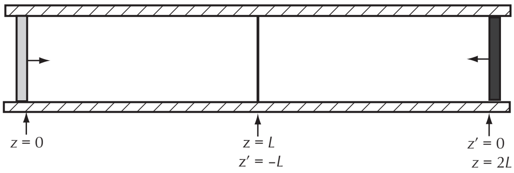
In the above equation A_{mn} are amplitude factors determined by the conditions at the exciting source. These amplitude factors differ depending on the values of m and n, namely on the particular mode involved. All of these non-planar modes have non-uniform pressures as well as polarities over the cross section of the waveguide. The next question should be, “Where is our familiar plane wave solution in all of this?” The answer again lies in a further examination of Fig. 16-17. Notice that when $m = 0$, the function J_0 has zero slope when $x = 0$, that is right at the origin. For this case, not only is $m = 0$ but n and k_{00} are zero as well while $J_0(k_{00}r)$ has the value of one independent of r. As $m = 0$, $\cos(m\theta)$ is unity independent of the angle θ and the solution becomes the familiar $p(z, t) = A\cos(\omega t - k_z z)$ with A being the pressure amplitude at the face of the piston that is uniform over the cross section of the tube. Furthermore, as k_{00} is zero, the phase velocity is a constant value c at all frequencies.

16.4 Plane Wave Tubes having Arbitrary Terminations

Next we turn our attention to tubes of finite length that are excited only with plane waves but have arbitrary terminations. A good starting point is a tube with length L excited by a piston as in our original case but which is terminated by a rigid barrier at its receiving end. This situation is depicted in the upper half of Fig. 16-18.



Actual Structure



Equivalent Structure

Figure 16-18. Tube terminated by a rigid barrier.

As the piston begins to move in the actual structure a plane wave propagates in the positive z-direction reaching the rigid barrier after a time lapse of L/c. As the barrier is rigid, there can be no particle displacement or velocity at the barrier at any time. The wave is reflected and then travels in the negative z-direction back to the source where it is again reflected but now by the moving piston. This process continues to repeat itself over and over and after numerous transits back and forth arrives at a steady state condition with the boundary conditions at the piston matching those of the piston's motion and those at the barrier corresponding to zero particle displacement as well as velocity. In the region $0 \le z \le L$ we ultimately have the superposition of two waves traveling in opposite directions. The wave equation is linear so the principle of superposition is applicable in arriving at a solution for a particular case. Furthermore, the solution to the wave equation that satisfies the given boundary conditions is unique so we can treat the problem by analyzing the equivalent structure in the lower half of Fig. 16-18 that is based on the method of images. In this technique the left (light colored) piston is the actual source and the right (dark colored) piston is the image source that is located just as far to the right of the barrier as the actual source is located to the left of the barrier. When the left piston displaces to the right, the right piston displaces similarly to the left. In the equivalent structure there is no barrier at $z = L$ although its position is indicated in the drawing. In the active interval $0 \le z \le L$ in which our solution will apply we sum the individual waves generated by the real source and the image source. When the left piston moves to the right the air in front of it is compressed so it produces a pressure wave given by $p_l = p_m \cos(\omega t - kz)$ where p_m is the pressure amplitude and k is the propagation constant along the z-axis. Here we have dropped the subscript on the propagation constant as we are dealing only with a plane wave. Similarly, when the image piston moves to the left it produces a pressure wave $p_r = p_m \cos(\omega t + kz')$. The situation with regard to the particle velocity is decidedly different however. The particle velocity wave generated by the left piston is $u_l = u_m \cos(\omega t - kz)$, however when the right piston moves to the left, the air particles are moving in the negative direction so $u_r = -u_m \cos(\omega t + kz')$. Furthermore, $z' = z - 2L$ so in the active interval $0 \le z \le L$ the total acoustic pressure is given by

$$p = p_m \cos(\omega t - kz) + p_m \cos(\omega t + kz - 2kL) \tag{16-51}$$

while the total particle velocity is given by

$$u = u_m \cos(\omega t - kz) - u_m \cos(\omega t + kz - 2kL) \tag{16-52}$$

These composite expressions describe standing waves. Our traveling waves in opposite directions have combined to form standing waves of both acoustic pressure and particle velocity. Now at the barrier where $z = L$, the acoustic pressure is $p = 2p_m \cos(\omega t - kL)$. This means that the pressure amplitude is doubled signifying that a normally incident pressure wave is reflected in phase at a rigid barrier. On the other hand, the particle velocity (u) at the barrier as given by $u = u_m \cos(\omega t - kL) - u_m \cos(\omega t - kL)$ is identically zero at all times, indicating that a normally incident particle velocity wave is reflected with a change of polarity or a phase shift of π radians at a rigid barrier. The important question is, "What are the conditions at the origin where the left piston is located?" Upon setting $z = 0$ in the general equations we find that the total acoustic pressure at the origin is now $p = p_m \cos(\omega t) + p_m \cos(\omega t - 2kL)$. Now it is important to remember at this point that all angles are expressed in radians. Suppose that the length L is exactly $\lambda/4$ at the operating frequency of the piston. Upon remembering that $k = (2\pi)/\lambda$, then $2kL = 2(2\pi/\lambda)(\lambda/4) = \pi$. When this is the case, the two pressure terms differ in phase by π and their sum is zero at all times t. This is a resonant condition. The acoustic pressure being identically zero means that the piston motion is completely unimpeded. This will be true also when L is any odd integral multiple of $\lambda/4$. This ideal is never exactly achieved in practice because there are always some very small viscous losses at the walls of the tube and the amplitude of the right piston motion is always slightly less than that of the left piston.

The conditions that exist at the exciting piston for a tube of arbitrary length L terminated in a rigid barrier are usually studied by means of the mechanical impedance presented to the piston as a result of its interaction with the air at the origin. This mechanical impedance is the ratio of the complex force acting on the air at the piston face to the complex particle velocity of the air at the piston face. The complex force is the acoustic pressure at the origin expressed as a complex exponential or phasor multiplied by the cross sectional area of the tube. The complex particle velocity is the complex exponential statement of the particle velocity also at the origin. Complex exponentials and phasors are described in detail in Chapter 6. The mechanical impedance then is calculated from

$$Z_m = \frac{p_m S}{u_m} \times \frac{e^{j\omega t} + e^{j(\omega t - 2kL)}}{e^{j\omega t} - e^{j(\omega t - 2kL)}} \tag{16-53}$$

This can be simplified by factoring and canceling common terms in both the numerator and denominator to yield,

$$\begin{aligned} Z_m &= \frac{p_m S}{u_m} \times \frac{1 + e^{-j2kL}}{1 - e^{-j2kL}} \\ &= \frac{p_m S}{u_m} \times \frac{e^{-j2kL}}{e^{-j2kL}} \times \frac{e^{jkL} + e^{-jkL}}{e^{jkL} - e^{-jkL}} \\ &= \frac{p_m S}{u_m} \times \frac{\cos(kL)}{j\sin(kL)} \\ &= -j\frac{p_m S}{u_m}\cot(kL) \end{aligned} \tag{16-54}$$

The conclusion is that the mechanical impedance presented to the piston is a negative reactance for positive values of the cotangent as indicated by the $-j$ in the final statement. If there were no viscous losses this impedance would be zero at the resonant condition where L equals odd integral multiples of $\lambda/4$ and would be infinite for the anti-resonant condition where L equals integral multiples of $\lambda/2$. The fact that the mechanical impedance is purely reactive means that once steady state conditions are reached the average power supplied by the piston becomes zero in the ideal case. It also means that the acoustic pressure and particle velocity in the standing wave differ in phase by $\pi/2$ radians or 90°.

Next we will explore the general technique that is applicable to plane wave tubes having arbitrary terminations including that of a rigid barrier, an open-ended tube, and any other given mechanical impedance. Finally, we will consider the interaction between a small loudspeaker employed to excite the tube and an improperly terminated plane wave tube.

As a reminder, the mechanical impedance is defined to be the ratio of the complex mechanical force applied to an object divided by the resulting complex mechanical velocity of the object

$$Z_m = \frac{F}{u} \tag{16-55}$$

Additionally, at any point in a sound field the ratio of the complex acoustic pressure to the resulting particle velocity is called the specific acoustic impedance at the point in question.

$$Z_s = \frac{p(z, t)}{u(z, t)} \tag{16-56}$$

Now consider a plane wave tube that is fitted with a piston at $z = 0$ and mechanical impedance of Z_L at the spatial point $z = L$. Let the piston displacement at any time be described by the phasor

$$\xi = \xi_m e^{j\omega t} \tag{16-57}$$

where,
ξ_m is the amplitude of the piston displacement,
ω is the angular frequency of piston oscillation.

Remember that the actual piston motion is given only by the real part of this phasor namely $\xi = \xi_m \cos(\omega t)$. In the general case Z_L does not properly terminate the tube so we must allow for both a primary wave and a reflected wave. In which case the phasor description of the two waves becomes

$$\xi(z, t) = Ae^{j(\omega t - kz)} + Be^{j(\omega t + kz)} \tag{16-58}$$

At the origin where $z = 0$, the boundary condition is satisfied by having this last expression match the given piston motion from which it is learned that $\xi_m = A + B$.

Another independent equation is required in order to determine A and B uniquely. This equation is obtained by recognizing that at $z = L$ the ratio of the acoustic force to the particle velocity at that point must be equal to the mechanical impedance at that point. It is necessary then to write the general expressions for the acoustic pressure and the particle velocity that are valid anywhere in the tube and particularly at $z = L$.

$$\begin{aligned} u(z,t) &= \frac{\partial \xi(z,t)}{\partial t} \\ &= j\omega\xi(z,t) \\ &= j\omega[Ae^{j(\omega t - kz)} + Be^{j(\omega t + kz)}] \end{aligned} \tag{16-59}$$

$$\begin{aligned} p(z,t) &= \rho_0 c^2 s(z,t) \\ &= -\rho_0 c^2 \frac{\partial \xi(z,t)}{\partial z} \\ &= \rho_0 c^2 (jk)[Ae^{j(\omega t - kz)} - Be^{j(\omega t + kz)}] \end{aligned} \tag{16-60}$$

where,

$s(z, t)$ is the condensation,

ρ_0 is the undisturbed air density.

Now the force at $z = L$ is the acoustic pressure at that point multiplied by the cross-sectional area, S, of the tube. Dividing the force by the particle velocity at $z = L$ leads to the second equation involving A and B.

$$Z_L = \rho_0 cS \frac{(Ae^{-jkL} - Be^{jkL})}{(Ae^{-jkL} + Be^{jkL})} \tag{16-61}$$

The two independent equations for A and B are now solved simultaneously to obtain

$$A = \frac{\xi_m}{2} \times \frac{\left[1 + \frac{Z_L}{\rho_0 cS}\right][1 + j\tan(kL)]}{1 + j\frac{Z_L}{\rho_0 cS}\tan(kL)} \tag{16-62}$$

$$B = \frac{\xi_m}{2} \times \frac{\left[1 - \frac{Z_L}{\rho_0 cS}\right][1 - j\tan(kL)]}{1 + j\frac{Z_L}{\rho_0 cS}\tan(kL)} \tag{16-63}$$

Knowing the values for A and B it is now possible to evaluate $\xi(z, t)$, $p(z, t)$, and $u(z, t)$ anywhere in the tube. In particular, at the input of the tube where z is zero we can determine the mechanical load or impedance that the tube presents to the motion of the piston. This term is called Z_0 and is calculated to be

$$\begin{aligned} Z_0 &= \frac{Sp(0,t)}{u(0,t)} \\ &= \frac{Z_L + j\rho_0 cS\tan(kL)}{1 + j\frac{Z_L}{\rho_0 cS}\tan(kL)} \end{aligned} \tag{16-64}$$

This last result is quite general and applies not only to tubes but other shapes as well as long as the operating wavelength is large compared with the largest dimension associated with the structure's cross-section. Our previous result for a tube terminated with a rigid barrier that was calculated earlier can be readily obtained from the general expression for Z_0 by dividing both numerator and denominator by Z_L and then allowing Z_L to approach infinity. Another observation with regard to this general case that is worthy of note is the behavior that occurs when the tube is driven at a frequency or frequencies such that the tube length is an integral number of half wavelengths. When this is true, the tangent terms in Z_0 are exactly zero and Z_0 becomes identically equal to Z_L and the tube's mechanical impedance opposing the piston's motion is the same as the mechanical impedance that terminates the tube. The half wavelength tube then acts as an ideal transformer having a turns ratio of 1:1.

Rather than being terminated in a rigid barrier or cap, suppose that the tube just ends abruptly at $z = L$ while being surrounded by a very large, ideally infinite, plane baffle. What is the terminating mechanical impedance in this instance? The answer is not zero because the air particles at the end of the tube must push against the outside air contained within a 2π solid angle when they suffer displacement by the forward traveling wave contained within the tube. In fact, the air particles at the end of the tube experience exactly the same impedance as that experienced by the front face of a piston that is radiating into a half space or 2π solid angle. Alternatively, the truncated end of the tube might just end in open space in which case the radiation is almost unconfined or experiences nearly a 4π solid angle. In the latter case the acoustic pressure is approximately one half of that of the former case. Since the force is directly proportional to the pressure, the impedance experienced by the truncated tube less the baffle is also approximately one-half that of the infinite baffle case. In either case, the terminating impedance is calculated through the employment of what is termed the piston impedance function. The piston impedance function has real and imaginary parts that are written as $R(2ka) + jX(2ka)$ where $k = (2\pi)/\lambda$ and a is the piston radius or, in this case, the inner radius of the tube. The real part of the piston impedance function can be expressed in terms of the first order Bessel function of the first kind that we encountered previously in this chapter while the imaginary part can be expressed in terms of the first order Struve function.

$$R(2ka) = 1 - \frac{J_1(2ka)}{ka} \tag{16-65}$$

$$X(2ka) = \frac{H_1(2ka)}{ka} \tag{16-66}$$

These functions are graphed in Fig. 16-19.

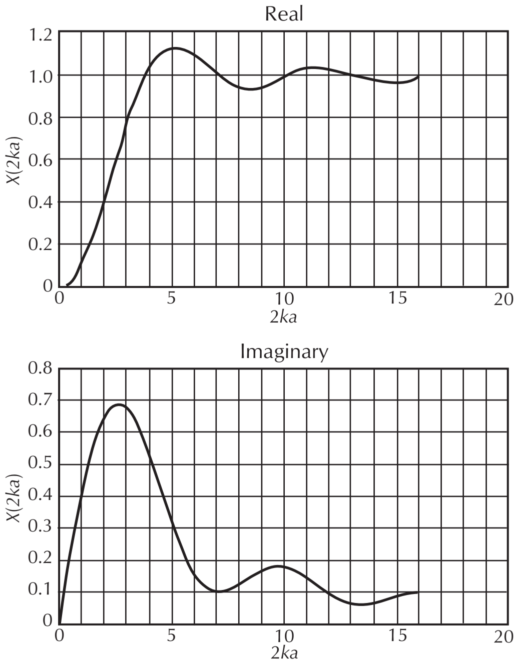


Figure 16-19. Real and Imaginary parts of the piston impedance function.

The terminating impedance of a truncated tube with a baffle expressed in terms of the piston impedance function is $Z_L = \rho_0 cS[R(2ka) + jX(2ka)]$ while that of the truncated tube without a baffle is approximately one half this amount.

It is very important to note that all of the foregoing takes no account of energy losses occurring within the air or at the interior surfaces of the tube. Air has both a viscous shear modulus that is a loss factor at the tube surfaces and a viscous bulk modulus that is a loss factor throughout the enclosed volume. Heat generation and conduction in the body of the gas and at the tube walls are even further considerations with regard to energy loss. A pursuit of these topics would carry us much further into the physics of fluids than we are prepared to go here. Even though the losses are small for short tubes of reasonable diameter, their inclusion significantly complicates the mathematics of the description of the process. The losses could be accounted for in our equations by allowing the propagation constant k to be complex with the form $\beta - j\alpha$. Replacing k in our equations by this complex form forces the particle displacement in our description to become

$$\xi = Ae^{-\alpha z}e^{j(\omega t - \beta z)} + Be^{+\alpha z}e^{j(\omega t + \beta z)} \tag{16-67}$$

In addition to having to redo the analysis employing this starting point, the problem is further complicated by the fact that both α and β are frequency dependent in a complicated fashion. The frequency dependence of β is particularly troublesome because the phase velocity being ω/β will no longer be independent of frequency. The problem can be handled exactly but the mathematics is more tedious. We will consider our results to be a first as well as useful approximation to the more exact ones.

Now we will use our approximate results to calculate the lowest resonant frequencies of a two-inch diameter loudspeaker whose front face is attached to a short tube of two inches inner diameter with the far end of the tube being terminated in a rigid cap. The back of the loudspeaker is enclosed by a small box. The air trapped in the box and the loudspeaker's suspension together act as a spring with a total stiffness of K. The suspension also furnishes a mechanical resistance R_m. The moving mass of the loudspeaker cone is M and the loudspeaker itself has a total mechanical impedance Z_{1s} with $Z_{1s} = R_m + j(\omega M - K/\omega)$. As we previously learned the closed tube loads the front face of the loudspeaker with a mechanical impedance that is $-j\rho_0 cS\cot(kL)$. The total mechanical impedance presented to the agency that drives the loudspeaker is then the sum of these two impedances with $Z_m = R_m + j(\omega M - K/\omega - \rho_0 cS\cot(kL))$. The driven loudspeaker will be at resonance for those frequencies where the total reactance in the mechanical impedance expression becomes zero or where $\omega M - K/\omega = \rho_0 cS\cot(kL)$. In this last equality, we replace k by ω/c and then replace ω by $2\pi f$ so that both sides can be plotted versus f. The intersection points of the two resulting curves identify the resonant frequencies. This was done by employing the parameters typical of a two-inch loudspeaker mounted on a two-inch tube of one-meter length. The results are presented in Fig. 16-20. For comparison purposes, the same calculation was performed taking account of air losses in the tube. These results are presented in Fig. 16-21.

The mechanical impedance at the input of a capped tube when the air losses in the tube are small is

$$Z_0 = \rho_0 c S \frac{\alpha L - j\cos(kL)\sin(kL)}{\sin^2(kL) + (\alpha L)^2\cos^2(kL)} \qquad (16\text{-}68)$$

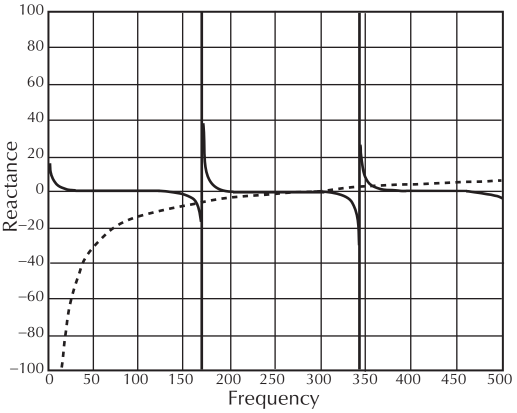


Figure 16-20. The solid curve is a plot of the cotangent reactance function with the vertical lines indicating the points of discontinuity of this function. The dashed curve is the loudspeaker reactance curve. Discounting the intersections at the discontinuous jumps, in the depicted frequency range the lowest resonance is 165 Hz, the middle resonance at 281 Hz is close to the loudspeaker's natural resonance without front loading of 290 Hz, and the upper resonance is at 362 Hz.

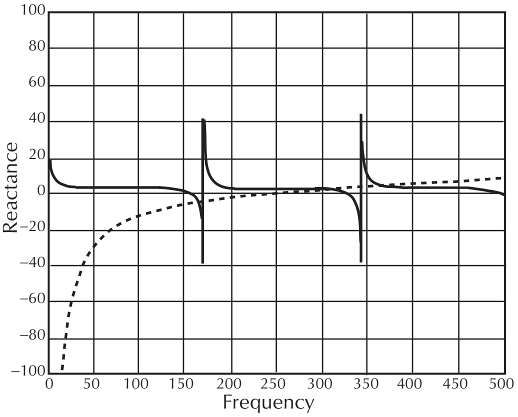


Figure 16-21. The solid curve is the reactance presented by the tube including the effect of air losses. The dashed curve is the loudspeaker reactance curve. The resonant frequencies indicated here are essentially the same as those of Fig. 16-20.

The real part of this expression is a mechanical resistance and along with the mechanical resistance of the loudspeaker broadens the shape of a resonance but does not effect its location. Setting α equal to zero reduces the impedance expression to that which was derived without considering air loss. There are resonances beyond the frequency range depicted, but the height of the dotted curve eventually exceeds that of the solid and there will be no further intersections beyond such a point.

16.5 Impedance Tube

The analyses discussed in the preceding section form the basis for a device employed in the determination of the mechanical impedance properties of materials that are often employed as acoustical energy absorbers. The structure of such a device is illustrated by Fig. 16-22.

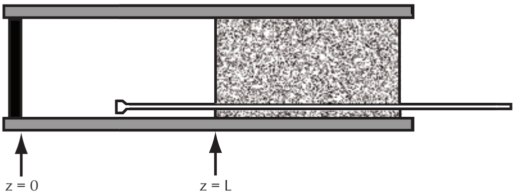


Figure 16-22. The illustration is an example of an impedance tube's basic structure.

An impedance tube consists of a hollow, rigid walled tube that is fitted with a driven piston whose right hand face has a rest position at $z = 0$. Beginning at $z = L$, the remainder of the volume of the tube is fitted with the sample of interest through which passes a thin, calibrated rod whose purpose is to control the position of a miniature pressure sensitive microphone. The piston undergoes an oscillatory motion with amplitude of 10^{-6} m at a frequency of 1000 Hz in the example to be presented here. In this example the static density of air, ρ_0, is 1.2 kg•m^{-3} and the sound speed, c, is 344 m•sec^{-1}. The wavelength of the plane wave generated by the piston's motion is 0.344 m and the distance L is adjusted to this value. The plane wave generated by the piston propagates toward the sample where a portion of the wave energy is reflected back towards the piston and the remainder is transmitted into the sample where it is subsequently absorbed assuming the sample is homogeneous and of sufficient length. Under steady state conditions the air-filled space in the interval $0 \le z \le L$ contains a plane wave traveling in the direction of increasing z and a wave of lower amplitude and shifted phase traveling in the direction of decreasing z. The superposition of these oppositely directed traveling waves produces a stationary or standing wave pattern in this space an example of which can be experimentally determined by positioning the movable miniature microphone at successive values of the coordinate z in the air filled space and measuring the acoustic pressure at each position. The results of such observations obtained with a particular sample are displayed in Fig. 16-23.

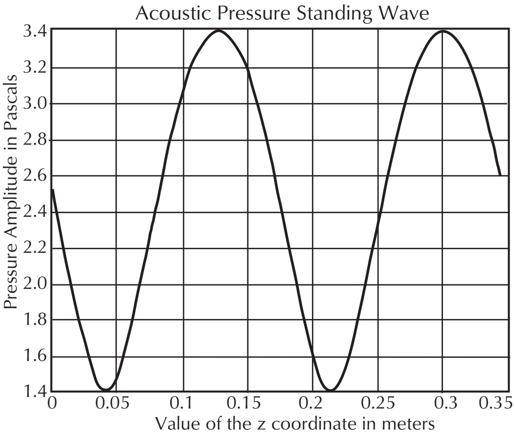


Figure 16-23. Pressure standing wave produced in an impedance tube.

The standing wave pattern of Fig. 16-23 features both minima and maxima. The minima are called nodes and the maxima are anti-nodes. Knowledge of the location of the first node and its associated pressure along with the pressure at the anti-node are sufficient experimental information to allow us to learn the mechanical impedance of the sample under test, as we will subsequently demonstrate. The first node is located at $z = 0.043$ m and is exactly one-eight of a wavelength from the origin. The acoustic pressure at this node has a value of 1.40 Pascals. The first anti-node has an acoustic pressure of 3.38 Pascals. The standing wave ratio or *SWR* is the quotient of these two values and is 2.414.

In our original discussion of a piston driven plane wave tube with an arbitrary termination we found the following two phasor equations describing the particle displacement and acoustic pressure present in the tube as well as the general expression for the terminating impedance at $z = L$.

$$\xi(z,t) = Ae^{j(\omega t - kz)} + Be^{j(\omega t + kz)} \tag{16-69}$$

$$p(z,t) = \rho_0 c^2 jk[Ae^{j(\omega t - kz)} - Be^{j(\omega t + kz)}] \tag{16-70}$$

$$Z_L = \rho_0 cS\frac{(Ae^{-jkL} - Be^{jkL})}{(Ae^{-jkL} + Be^{jkL})} \tag{16-71}$$

Identities for these three equations in a more convenient form for the present purpose appear with the following structure.

$$\xi(z,t) = A\left[1 + \frac{B}{A}e^{j2kz}\right]e^{j(\omega t - kz)} \tag{16-72}$$

$$p(z,t) = \rho_0 c^2 jkA\left[1 - \frac{B}{A}e^{j2kz}\right]e^{j(\omega t - kz)} \tag{16-73}$$

$$Z_L = \rho_0 cS\frac{\left(1 - \frac{B}{A}e^{j2kL}\right)}{\left(1 + \frac{B}{A}e^{j2kL}\right)} \tag{16-74}$$

Recall that A and B are in general complex amplitude factors and that the magnitude of B is less than that of A. We will then represent their quotient as $Fe^{j\theta}$, where F is less than one and is the magnitude of B divided by the magnitude of A and the angle theta is the angle of B minus the angle of A. Making use of this form of the quotient, the equation for the acoustic pressure takes on the following form.

$$p(z,t) = \rho_0 c^2 jkA[1 - Fe^{j(\theta + 2kz)}]e^{j(\omega t - kz)} \tag{16-75}$$

The pressure amplitude will have its first minimum when the term in the bracket first acquires a minimum value. This minimum is simply $1 - F$. This occurs when $\theta + 2kz$ first becomes equal to zero because Euler's theorem states that $e^{j(0)} = \cos(0) + j\sin(0) = 1$. Recall that $k = (2\pi/\lambda)$ and that the first node in the pressure curve occurs at $z = \lambda/8$ so

$$\theta + 2\frac{2\pi\lambda}{\lambda\ 8} = 0,\ \therefore \theta = -\frac{\pi}{2} \tag{16-76}$$

At the pressure node the value of the exponential term multiplying F is just one so that the minimum pressure is proportional to $1 - F$. By the same token, the pressure at the anti-node occurs when the value of the exponential term multiplying F is minus one and the maximum pressure is proportional to $1 + F$. The standing wave ratio or *SWR* was experimentally found to be 2.414. Therefore we may write

$$\frac{1+F}{1-F} = 2.414\text{, from which, } F = 0.4142 \tag{16-77}$$

Finally, we are now in position to determine the value of the terminating mechanical impedance Z_L. We purposely chose the length L to be a wavelength at the operating frequency. This choice makes the angle in the exponential factors of the equation for Z_L have the value of either $+4\pi$ or -4π for which the cosine of the angle is one and the sine of the angle is zero in either case, so Z_L is simplified to

$$Z_L = \rho_0 cS \frac{\left[1 - Fe^{j\left(-\frac{\pi}{2}\right)}\right]}{\left[1 + Fe^{j\left(-\frac{\pi}{2}\right)}\right]} \qquad (16\text{-}78)$$

Upon substituting the magnitude of $F = 0.4142$ in the above equation we arrive with the result

$$Z_L = \rho_0 cSe^{j\left(\frac{\pi}{4}\right)} \qquad (16\text{-}79)$$

S is just the cross-sectional area of the impedance tube and if we divide the mechanical impedance of the terminating medium by S we will be left with the specific acoustic impedance of the medium in this instance, $\rho_0 ce^{j(\pi/4)}$.

An acoustic pressure wave propagating in a medium having a complex specific acoustic impedance as in the above example loses energy as both a function of the distance traveled as well as the frequency and has a phase velocity that is frequency dependent.

16.6 More General Waves

The foregoing might leave the false impression that plane waves exist only in pipes, tubes, or other bounded structures. This is far from the truth. Plane waves can readily exist in unbounded regions. In fact, acoustic radiation from any source of finite dimensions will eventually approach plane wave status at sufficiently large distances from the source. We will observe this effect in the next section dealing with simple spherical waves.

Pulsating Sphere

If we were asked to visualize the simplest possible structure that could generate sound waves in three dimensions based only on what we have learned so far along with all our experiences from childhood onward we probably would imagine a pulsating, impermeable spherical surface. Such a surface might have a nominal radius, a, and would be alternately expanding and contracting with all motion being along radial lines. No loudspeaker has ever been constructed in exactly this way though if one did exist it would have the useful property of radiating uniformly in all directions. Imagine that the sphere is pulsating at a fixed frequency f and that the radial velocity at any point on the surface of the sphere is given by the real part of $u_m e^{j(\omega t)}$ where $\omega = 2\pi f$ and u_m is the amplitude of the radial surface velocity. Under these conditions the alternate small expansions and contractions of the sphere will be radiating an acoustic signal at the frequency f that will have some constant average power. As this signal diverges uniformly into the space surrounding the pulsating sphere, this power will be spreading out and flowing through concentric surfaces having progressively increasing radii. The amount of power is constant but the area through which it must flow is growing at a rate that is proportional to the radial distance from the center of the pulsating sphere quantity squared. As a result, the intensity of the radiated acoustic signal is decreasing at a rate that is inversely proportional to the square of the distance from the origin at the center of the pulsating sphere. The intensity in general is proportional to the square of the acoustic pressure. Therefore, the acoustic pressure amplitude must be decreasing inversely with the first power of the radial distance from the origin. In other words, the pressure amplitude at any distance r when multiplied by r is a constant for any value of r equal to or greater than the nominal radius of the sphere. In the case of plane waves it was found that the pressure amplitude was constant independent of the distance from the origin and that the wave equation governing the acoustic pressure was

$$\frac{\partial^2 p}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} \qquad (16\text{-}80)$$

In the spherical wave case under examination, the product of the pressure with radial distance from the origin is a constant independent of the choice of distance from the origin so we propose as a wave equation

$$\frac{\partial^2 (pr)}{\partial r^2} = \frac{1}{c^2}\frac{\partial^2 (pr)}{\partial t^2} \qquad (16\text{-}81)$$

where,

r is the radial distance from the center of the pulsating sphere

$r \geq a$, where a is the nominal radius of the pulsating sphere.

A solution to the proposed wave equation can be obtained through the same considerations that were made in the plane wave case if one recognizes that the dependent variable is now pr rather than just p itself thus,

$$p = \frac{A}{r}e^{j(\omega t - kr)} \qquad (16\text{-}82)$$

where,

A is a complex amplitude factor that must be determined by the acoustic pressure at the surface of the pulsating sphere where r has the value a. At this point, we only know the radial velocity at the surface of the sphere. It is necessary then to explore the relationship between the acoustic pressure and the particle velocity in the spherical wave case. When Newton's second law was applied in the derivation of the plane wave case it was learned that

$$-\frac{\partial p}{\partial z} = \rho_0 \frac{\partial u}{\partial t} \tag{16-83}$$

A similar application to the present case yields

$$-\frac{\partial p}{\partial r} = \rho_0 \frac{\partial u}{\partial t} \tag{16-84}$$

Substitution of Eq. 16-82 into the left side of Eq. 16-84 yields

$$\begin{aligned}\left(\frac{1}{r} + jk\right)p &= \rho_0 \frac{\partial u}{\partial t} \\ &= \rho_0 j\omega u\end{aligned} \tag{16-85}$$

Upon solving Eq. 16-85 for the ratio of the acoustic pressure to the particle velocity one obtains a quantity that is called the specific acoustic impedance of air for spherical waves.

$$\begin{aligned}Z_s &= \frac{\rho_0 ckr}{\sqrt{1 + k^2r^2}} e^{j\left(\arctan\frac{1}{kr}\right)} \\ &= (\rho_0 c)\cos(\theta)e^{j\theta}\end{aligned} \tag{16-86}$$

where,

the angle θ as depicted in Fig. 16-24 represents the phase difference between the acoustic pressure and the particle velocity at the radial distance r from the source,

$(\rho_0 c)\cos(\theta)$ is the pressure amplitude divided by the particle velocity amplitude at the same radial distance from the source.

Now $p = Z_s u$ which means that at the surface of the sphere where $r = a$,

$$\frac{A}{a} e^{j(\omega t - ka)} = \frac{\rho_0 cka}{\sqrt{1 + k^2a^2}} u_m e^{j\left(\omega t + \arctan\frac{1}{ka}\right)} \tag{16-87}$$

When Eq. 16-87 is solved, A is found to be

$$A = \frac{\rho_0 cka^2}{\sqrt{1 + k^2a^2}} u_m e^{j\left(ka + \arctan\frac{1}{ka}\right)} \tag{16-88}$$

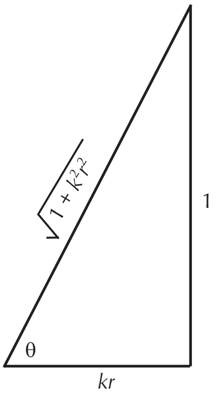


Figure 16-24. The angle θ is the angle of the specific acoustic impedance of air for spherical waves.

When this value of A is substituted into Eq. 16-82 the result becomes

$$p = \frac{\rho_0 cka^2}{r\sqrt{1 + k^2a^2}} u_m e^{j\left(\omega t - kr + ka + \arctan\frac{1}{ka}\right)} \tag{16-89}$$

Such a formidable equation might evoke thoughts of a career change. Fortunately, matters can be considerably simplified by restricting attention to pulsating spheres which are small such that $ka \langle\langle 1$. It is desirable to do this particularly when this theory is to be applied to sound sources that are other than of spherical shape. For $ka \langle\langle 1$,

$$A \approx \rho_0 cka^2 u_m e^{j\left(\frac{\pi}{2}\right)} = j\rho_0 \omega a^2 u_m \tag{16-90}$$

When this approximate expression for A is substituted into Eq. 16-82 the pressure becomes

$$p = \frac{j\rho_0 \omega a^2}{r} u_m e^{j(\omega t - kr)} \tag{16-91}$$

Finally, it is useful to introduce the area of the radiating source that for a sphere of radius a is $S = 4\pi a^2$ so that Eq. 16-91 becomes

$$p = \frac{j\rho_0 \omega S}{4\pi r} u_m e^{j(\omega t - kr)} \tag{16-92}$$

A few extra words might be of value here with regard to the mathematical technique employed in the above analysis dealing with a real physical variable such as acoustic pressure or particle velocity. In physical situations where the time t enters the picture in the form of a harmonic function such as $\cos(2\pi ft \pm \alpha)$ or $\cos(\omega t \pm \alpha)$, the mathematical operations of differentiation and integration with respect to time are greatly expedited when the harmonic function is written in the complex exponential form. The angle α may be any angle or collection of angles that are not time dependent though they may be frequency dependent. After the necessary operations have been performed, the actual real physical quantity is obtained by taking only the real part of the final complex exponential expression. An example is perhaps in order. Recall Euler's theorem that $e^{j(\beta)} = \cos(\beta) + j\sin(\beta)$ where β is any angle, so we write the leading j in Eq. 16-92 as $e^{j(\pi/2)}$ and rewrite the equation as

$$p = \frac{\rho_0 S u_m}{4\pi r} e^{j\left(\omega t - kr + \frac{\pi}{2}\right)} \tag{16-93}$$

Next we expand Eq. 16-93 according to Euler to obtain separated real and imaginary parts.

$$\begin{aligned} p &= \frac{\rho_0 \omega S u_m}{4\pi r}\cos\left(\omega t - kr + \frac{\pi}{2}\right) \\ &\quad + j\frac{\rho_0 \omega S u_m}{4\pi r}\sin\left(\omega t - kr + \frac{\pi}{2}\right) \end{aligned} \tag{16-94}$$

Finally, it is only the real part of Eq. 16-94 that describes the actual acoustic pressure as a function of position and time so

$$\begin{aligned} p(r, t) &= \frac{\rho_0 \omega S u_m}{4\pi r}\cos\left(\omega t - kr + \frac{\pi}{2}\right) \\ &= p_m \cos\left(\omega t - kr + \frac{\pi}{2}\right) \end{aligned} \tag{16-95}$$

where,

p_m is the acoustic pressure amplitude,

ρ_0 is the undisturbed air density,

ω is the angular frequency,

S is the surface area of the radiating source,

u_m is the amplitude of the surface velocity of the source,

r is the radial distance from the source.

Now that the acoustic pressure is known it becomes possible to calculate the particle velocity of the air. The specific acoustic impedance of air for spherical waves is inherently a complex quantity so we first represent both Z_s as well as $p(r,t)$ as complex exponentials and make use of the fact that

$$\begin{aligned} u_a(r, t) &= \frac{p(r, t)}{Z_s} \\ &= \frac{\dfrac{\rho_0 \omega S u_m}{4\pi r} e^{j\left(\omega t - kr + \frac{\pi}{2}\right)}}{\rho_0 c \dfrac{kr}{\sqrt{1 + k^2 r^2}} e^{j\theta}} \\ &= \frac{p_m}{(\rho_0 c)\cos(\theta)} e^{j\left(\omega t - kr + \frac{\pi}{2} - \theta\right)} \end{aligned} \tag{16-96}$$

where,

$u_a(r,t)$ is the particle velocity of the air at the distance r,

p_m is the acoustic pressure amplitude at the distance r,

$(\rho_0 c)\cos(\theta)$ is the magnitude of the specific acoustic impedance of air at the distance r.

The physical value for the actual particle velocity is just the real part of the result of Eq. 16-96 and is given by

$$u_a(r, t) = \frac{p_m}{(\rho_0 c)\cos(\theta)}\cos\left(\omega t - kr + \frac{\pi}{2} - \theta\right) \tag{16-97}$$

Two items are worthy of note in examining Eq. 16-97. Firstly, the amplitude of the particle velocity is the amplitude of the acoustic pressure divided by the magnitude of the specific acoustic impedance and secondly, the phase difference between the acoustic pressure and the particle velocity is the angle of the specific acoustic impedance as displayed in Fig. 16-24. When we expand Eq. 16-86, that is the definition of the specific acoustic impedance of air for spherical waves, into its separated real and imaginary parts we obtain

$$Z_s = \frac{\rho_0 c k^2 r^2}{1 + k^2 r^2} + j\frac{\rho_0 c k r}{1 + k^2 r^2} \tag{16-98}$$

As an observer approaches more and more distant points from a simple spherical source or monopole as it is called, the term $k^2 r^2$ grows larger and larger as compared with one and the real part of Eq. 16-98 approaches the value $\rho_0 c$. Additionally, in these distant regions the imaginary part of Z_s is getting smaller and smaller and is approaching zero with the result that Z_s is approaching the value that is characteristic of plane waves in air where the acoustic pressure and particle velocity are in phase. So at very large distances from a simple source plane and spherical waves are indistinguishable by only local measurements.

16.7 Acoustic Intensity

Acoustic intensity is a statement of the value of the acoustic energy flow per unit area per unit of time. This is the same as acoustic power per unit of area. We measure and speak of both instantaneous intensity and average intensity. The instantaneous intensity is the product of the acoustic pressure with the particle velocity at any given instant and at the same point in space as expressed in the following statement.

$$I(t) = p(t)u(t) \tag{16-99}$$

$I(t)$ is a vector quantity whose direction at any moment is that of the instantaneous transmission of acoustic power. The acoustic pressure is a scalar quantity that has no direction in space although it does possess an algebraic sign that may change from moment to moment. The vector character of the intensity arises from the particle velocity that is inherently a directed quantity. The average intensity over a specified time interval is the acoustic energy per unit area transmitted through an area perpendicular to the direction of wave propagation in the specified time interval averaged over the specified time interval. Mathematically this amounts to

$$<I(t)> = \frac{1}{T}\int_0^T p(t)u(t)dt \tag{16-100}$$

In the case of harmonic plane waves, the acoustic pressure and particle velocity are in phase while the amplitude of the particle velocity is that of the pressure divided by $\rho_0 c$. In this instance Eq. 16-100 becomes

$$\begin{aligned} <I(t)> &= \frac{1}{T}\int_0^T \frac{p_m^2}{\rho_0 c}\cos^2(\omega t - kz)dt \\ &= \frac{1}{2}\frac{p_m^2}{\rho_0 c} \end{aligned} \tag{16-101}$$

where,

T is the period or an integral number of periods corresponding to the particular frequency involved.

In the case of sinusoidal functions of time, which is true of Eq. 16-101, one may replace the pressure amplitude by the root mean square value of the pressure to obtain the simpler result $<I(t)> = p_{rms}^2/\rho_0 c$. This is possible for sinusoids because the root mean square value of such functions is the amplitude divided by the square root of two. I hasten to add, however, that this is not true of time functions in general.

In the case of spherical harmonic waves from a simple source, the calculation is a little more involved as the acoustic pressure and particle velocity may well differ in phase. This is similar to the electrical case in ac circuits where the sinusoidal voltage and current are not necessarily in phase. The phase difference between the voltage and the current is compensated for by introducing the power factor into the calculation of the average power. The power factor in the electrical case is the cosine of the angle of the electrical impedance where as in our case it is the cosine of the angle that represents the phase difference between acoustic pressure and particle velocity, i.e., the angle of the specific acoustic impedance. The formal statement of the average intensity appears as

$$\begin{aligned} &<I(t)> = \\ &\frac{1}{T}\int_0^T p_m\cos\left(\omega t - kr + \frac{\pi}{2}\right)\frac{p_m}{(\rho_0 c)\cos(\theta)}\cos\left(\omega t - kr + \frac{\pi}{2} - \theta\right)dt \end{aligned} \tag{16-102}$$

In the process of performing the integration in Eq. 16-102 use is made of a trigonometric identity in which $\cos(A-B) = \cos(A)\cos(B) + \sin(A)\sin(B)$ where we identify $\omega t - kr + \pi/2$ as A and θ as B. We apply this identity to the second cosine function in the integral. Next, we recognize that the sine terms in the expansion will not make any contribution to the integral over a complete period of the signal so the integral now appears in the following form

$$\begin{aligned} &<I(t)> = \\ &\frac{1}{T}\int_0^T \frac{p_m^2}{(\rho_0 c)\cos(\theta)}\cos^2\left(\omega t - kr + \frac{\pi}{2}\right)\cos(\theta)dt \end{aligned} \tag{16-103}$$

The final result is identical to that of the plane wave case so for both cases

$$\begin{aligned} <I(t)> &= \frac{1}{2}\frac{p_m^2}{\rho_0 c} \\ &= \frac{p_{rms}^2}{\rho_0 c} \end{aligned} \tag{16-104}$$

All of the foregoing analysis with regard to a simple spherical source has applications to sources of other shapes as long as the largest nominal dimension of the source, say a, is such that $ka \langle\langle 1$ and all portions of the surface of the source displace in phase.

16.8 Boundaries

The simple spherical source in all of the preceding analysis has been assumed to be completely isolated from other sources as well as any environmental boundaries. Next we need to consider what happens when the simple source is placed very close to a large rigid barrier such as indicated in the simple profile sketch of Fig. 16-25.

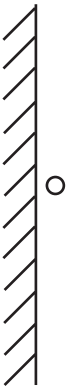


Figure 16-25. A simple spherical source located very close to an ideally infinite rigid barrier.

We have learned that all air particle motion must occur along radial lines directed away from the center of the source. At the barrier itself however, there can be no air particle motion perpendicular to the barrier itself because of its rigidity while there is no restriction on air particle motion tangential to the barrier surface. In the language of mathematics, the normal component of the air particle velocity must vanish at the barrier. For all of the space to the right of the barrier, doing two things can readily satisfy this condition. First we replace the barrier by an imaginary plane surface with the real source located to the right of the plane by exactly the same distance as it was from the actual barrier. Next we consider an image source with identical properties to that of the real source placed just as far to the left of the plane as the real source is to the right. This new situation is depicted in Fig. 16-26 where the distance between the source and the plane has been exaggerated for clarity.

In the three dimensional space to the right of the infinite plane defined by $x = 0$ the acoustic pressure and particle velocity are exactly described by the actual simple spherical source along with its image. We illustrate this by making calculations at three points. In the figure the z-axis is pointed towards the reader. At the origin x, y, and of course z are each equal to zero. The real source is located at $(\delta, 0, 0)$ while the image source is located at $(-\delta, 0, 0)$. At the origin the particle velocities of the real and image sources are of the same magnitude, oppositely directed, and both perpendicular to the plane. Their vector sum is of course zero thus satisfying the required boundary condition. Specifying a z value other than zero will produce the same change in the magnitudes of the particle velocities of the two sources while adding an equal tangential component to each. The normal components of each are still equal and oppositely directed so that the boundary condition remains satisfied. At the point $(0, y, 0)$ the radial lines drawn from each source intercept the plane such that the x components of particle velocity along the lines are oppositely directed thus canceling at the plane while the y components are in the same direction and thus add to that double what a single source alone could produce. Finally at the point $(x, 0, 0)$ we may deal with the acoustic pressure directly as there is no boundary condition to be dealt with. The acoustic pressure amplitude at this point produced by the actual source is $\rho_0 \omega S u_m / 4\pi(x-\delta)$ while that produced by the image source is $\rho_0 \omega S u_m / 4\pi(x+\delta)$. Now if δ is very small compared with x meaning that the actual source is placed very close to the barrier then the combination of these two pressure terms yields the result $p_m = \rho_0 \omega S u_m / 2\pi x$ a value that is twice that of a single source without a barrier. Finally, consider a general point with coordinates (x, y, z). For the actual source we would have a pressure amplitude of $\rho_0 \omega S u_m / 4\pi\sqrt{(x-\delta)^2 + y^2 + z^2}$ while for the image source the amplitude would be $\rho_0 \omega S u_m / 4\pi\sqrt{(x+\delta)^2 + y^2 + z^2}$.

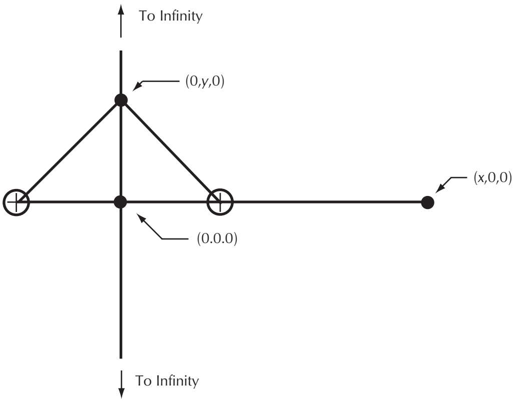


Figure 16-26. The method of real source and image as applied to a single source with nearby infinite rigid barrier.

Again if the actual source is very close to the barrier such that δ is very small compared with x then the combined pressure amplitude is double that produced by the actual source without the presence of a barrier. This result is termed half space radiation from a single source as expressed by

$$p(r, t) = \frac{\rho_0 \omega S u_m}{2\pi r}\cos\left(\omega t - kr + \frac{\pi}{2}\right) \qquad (16\text{-}105)$$

where,

$$r = \sqrt{x^2 + y^2 + z^2}$$

The method of images can be applied in any circumstance when a barrier is present even when no restriction is placed on the actual source's location. One simply defines separate radial distances from the actual and image source and uses these different values of radial distance in both the amplitude calculation as well as the phase calculation. For example, suppose the actual source has the coordinates (x_a, 0, 0) where x_a is any positive value. The radial distance from the actual source to any space point to the right of the barrier is then $r_a = \sqrt{(x - x_a)^2 + y^2 + z^2}$ while the radial distance from the image source is $r_i = \sqrt{(x + x_a)^2 + y^2 + z^2}$. The acoustic pressure at the general space point (x, y, z) to the right of the barrier is then

$$p(x, y, z, t) = \frac{\rho_0 \omega S u_m}{4\pi r_a}\cos\left(\omega t - kr_a + \frac{\pi}{2}\right) + \frac{\rho_0 \omega S u_m}{4\pi r_i}\cos\left(\omega t - kr_i + \frac{\pi}{2}\right) \qquad (16\text{-}106)$$

A cone type loudspeaker mounted in a back-enclosed box far from any boundaries will act as a simple spherical source radiating into all of space at low frequencies where the wavelengths are large compared with the dimensions of the enclosure. Similarly, such a loudspeaker placed on the floor in the middle of a large room will act as a simple spherical source radiating into a half space again only at low frequencies.

16.9 Acoustic Dipole

An acoustic dipole consists of two simple spherical sources separated by a small distance d and whose properties are identical with the exception that their surface velocities differ by pi radians or 180°. Such an arrangement is depicted in Fig. 16-27.

Fig. 16-27 is a sketch of an acoustic dipole located in the yz plane. The x- axis is towards the reader. The radial lines converge on a general space point. The radial distance from the origin is represented by r, while the radial distance from the expanding or positive source is r_1, and that from the contracting or negative source is r_2. These distances are all related in that:

$$r_1 = \sqrt{r^2 - dr\cos(\theta) + (d/2)^2}$$

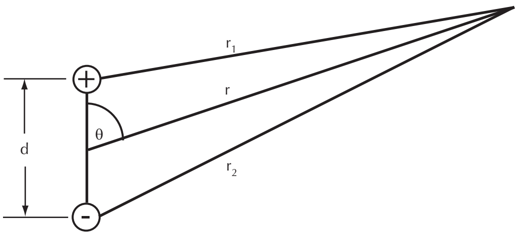


Figure 16-27. The out of polarity simple sources are located at $z = d/2$ and $z = -d/2$.

and

$$r_2 = \sqrt{r^2 + dr\cos(\theta) + (d/2)^2}.$$

The general expression for the acoustic pressure produced by the dipole without any approximations would have the form

$$p = \frac{\rho_0 \omega S u_m}{4\pi r_1}\left[\cos\left(\omega t - kr_1 + \frac{\pi}{2}\right)\right] - \frac{\rho_0 \omega S u_m}{4\pi r_2}\left[\cos\left(\omega t - kr_2 + \frac{\pi}{2}\right)\right] \qquad (16\text{-}107)$$

Viewing what has been developed so far, some general observations can be made without having to solve any equations. When the observation point is such that the angle theta is 90° then the observation point is equidistant from both sources and the acoustic pressure is zero. When the angle theta is less than 90° the positive source is dominant as the observation point is always closer to the positive source while for theta greater than 90° the negative source becomes dominant with r_2 being less than r_1. If we sacrifice total generality by placing some restrictions on the size of the source separation, that is, the size of d, we can make some useful simplifying approximations that lead to interesting conclusions with regard to the behavior of the acoustic pressure amplitude. When it is required that the separation between the two sources, d, be much less than the radial distance, r, as well as the wavelength, λ, then the following approximations are quite accurate.

$$r_1 \approx r\left(1 - \frac{1}{2}\frac{d}{r}\cos(\theta)\right)$$
$$r_2 \approx r\left(1 + \frac{1}{2}\frac{d}{r}\cos(\theta)\right) \qquad (16\text{-}108)$$

When the approximations expressed in Eq. 16-108 are substituted into the general Eq. 16-107 the resulting equation can be simplified to show that the combined acoustic pressure amplitude produced by the dipole can be found from the relatively simple expression

$$p_m = \frac{\rho_0\omega^2 S u_m}{4\pi c r}(d\cos(\theta)) \tag{16-109}$$

where,
c is the sound phase velocity,
u_m is the amplitude of the surface velocity of each source,
ω is the angular frequency,
r is the radial distance from the center of the dipole.

It should be apparent from Eq. 16-108 that the dipole is a poor radiator particularly at low frequencies as the pressure amplitude has a quadratic dependence on the angular frequency. This amplitude expression has been written so as to also exhibit the change in polarity that occurs when theta exceeds 90°. Fig. 16-28 displays a polar plot of the normalized directivity of an acoustic dipole. The axis of the dipole is along the horizontal axis.

An obvious example of an acoustic dipole is a bare frame cone type loudspeaker where the separation distance, d, is the actual loudspeaker frame diameter. When such a loudspeaker is mounted at the center of a large flat baffle the source separation distance then becomes the diameter of the baffle and the loudspeaker radiates more efficiently at all frequencies because of the enlarged value of d.

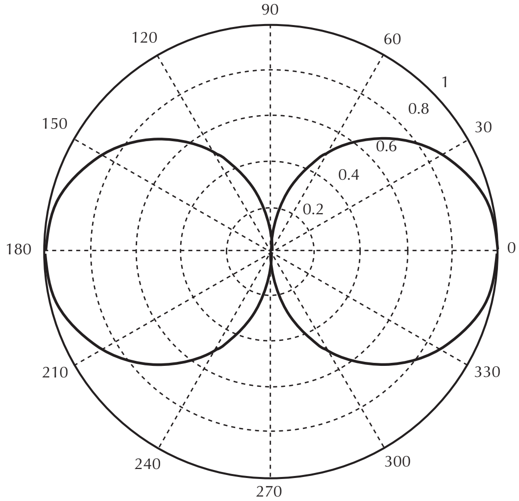


Figure 16-28. Polar plot of the normalized directivity of an acoustic dipole.

Bibliography

Lawrence E. Kinsler, Austin Frey, et. al. *Fundamentals of Acoustics,* 4th ed. New York: John Wiley and Sons, 2000.

Allan D. Pierce. *Acoustics: An Introduction to Physical Principles and Applications,* New York: Mc-Graw Hill, 1981

Chapter 17

Microphones

by Eugene Patronis, Jr.

17.1 The Microphone as the System Input 315
17.2 Microphone Sensitivity 315
17.3 Thermal Noise 317
17.4 Microphone Selection 324
  Carbon 324
  Capacitor 324
  Moving Coil 324
  Ribbon 325
  Piezoelectric 325
  Matching Talker to Microphone 325
17.5 Nature of Response and Directional Characteristics 325
17.6 Boundary Microphones 330
17.7 Wireless Microphones 335
17.8 Microphone Connectors, Cables, and Phantom Power 339
17.9 Measurement Microphones 341
  Measurement Microphone Types 341
17.10 Microphone Calibrator 342
Bibliography 344

Microphones are the input components of an extended voice reinforcement sound system. In earlier chapters it has been established that there is a relationship between the choice of type and location of microphone, loudspeaker, talker, and listener that affects achievable acoustic gain and intelligibility. The next concern is what happens to the signal from the acoustic input, into the microphone, through the electrical gains and losses of the sound system components, until sufficient electrical power reaches the loudspeaker or loudspeaker array.

17.1 The Microphone as the System Input

There are many excellent texts on how the various microphone types transduce acoustic pressure into an electrical signal. Such texts also explore the particular microphone structures necessary to achieve particular polar response patterns.

Microphones can be divided into several categories:

1. Measurement.
2. Entertainment.
3. Reinforcement.
4. Broadcast.
5. Recording.

Each category has its own special characteristics. For example, an entertainment microphone may feature the proximity effect to add body to a thin voice and the presence effect to add sparkle. On the other hand, measurement microphones feature a wide-band uniform pressure response. It is a fact that visual appearance, tactile qualities, and certain accessory convenience features in addition to technical characteristics play major roles in the professional choice of a microphone. The following parameters are of concern to sound system engineers:

1. Sensitivity.
2. Polar response.
3. Amplitude response.
4. Impedance rating.
5. Polarity.
6. Phase response.
7. Maximum acoustical input level.
8. Distortion properties.
9. Special features such as noise cancellation, wireless, and internal filters.

Knowledge of certain technical parameters is required in order to integrate a commercial microphone into a workable sound system. When reliable technical data is not at hand it becomes necessary to make field measurements to gain the information required.

17.2 Microphone Sensitivity

In order to determine the electrical input level to a sound system we need to measure the electrical output generated by the system microphone when it is subjected to a known sound pressure (*SP*). In making such measurements an L_P of 94 dB (1 Pa) is recommended as this value is well above the normally encountered ambient noise levels.

Everyone seriously interested in the field of professional sound should own or have easy access to a precision sound level meter (SLM). Among other uses, a SLM is required to measure ambient noise, to calibrate sources, and on occasion to serve as input for frequency response, reverberation time, signal delay, distortion, and acoustic gain measurements.

Setting up the microphone measurement system shown in Fig. 17-1 requires a pink noise generator, a micro-voltmeter, a high-pass and low-pass filter set such as the one illustrated in Fig. 17-2, a power amplifier, and a well-constructed test loudspeaker in addition to the SLM.

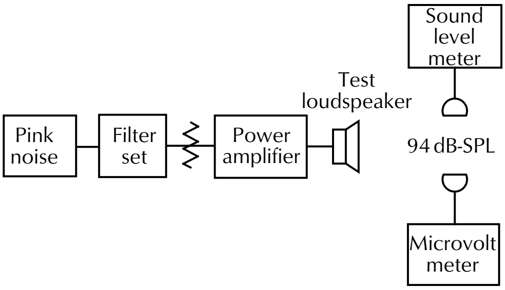


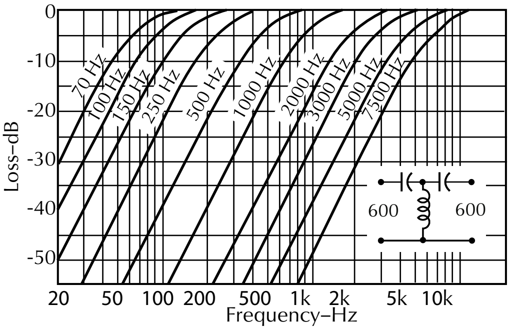
Figure 17-1. Measuring microphone sensitivity.

Select a measuring point (about 5 ft to 6 ft) in front of the loudspeaker, and place the SLM there. Adjust the system until the SLM reads an L_P of 94 dB (a band of pink noise from 250–5000 Hz is excellent for this purpose). Now substitute the microphone to be tested for the SLM. Take the microphone open circuit voltage reading on the micro-voltmeter. The voltage sensitivity of the microphone can then be defined as

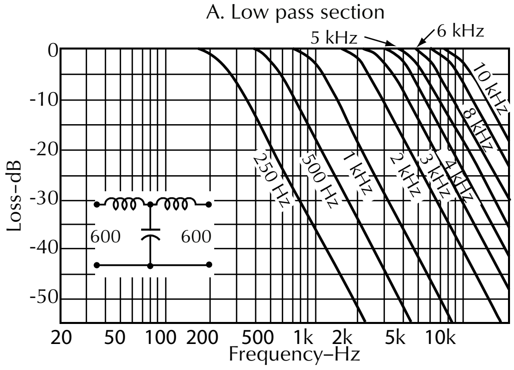
$$S_V = 20\ \mathrm{dB}\log(E_o) \tag{17-1}$$

where,

S_V is the voltage sensitivity expressed in decibels referenced to 1 volt for a 1 Pa acoustic input to the microphone,



A. Low pass section



B. High pass section

Figure 17-2. Response characteristics of a passive filter set. (*Courtesy United Recording Electronics Industries.*)

E_o is the open circuit output of the microphone in volts.

The open circuit voltage output of the microphone when exposed to some other arbitrary acoustic level L_P is calculated from

$$E_o = 10^{\left(\frac{S_V + L_P - 94}{20}\right)} \tag{17-2}$$

where,

E_o is now the open circuit voltage output of the microphone for an arbitrary acoustic input of level L_P.

For example suppose a sample microphone is tested by the conditions of Fig. 17-1 with the result that the open circuit voltage is found to be 0.001 V. The voltage sensitivity of this microphone as calculated from Eq. 17-1 is then

$$\begin{aligned} S_V &= 20\ \text{dB}\log(0.001) \\ &= -60\ \text{dB} \end{aligned}$$

This result would be read as –60 dB referenced to 0 dB being 1 volt per pascal (1 V/Pa). If this same microphone were exposed to an acoustic input level of 100 dB rather than the test value of 94 dB, then its open circuit output voltage from Eq. 17-2 would become

$$\begin{aligned} E_o &= 10^{\left(\frac{-60+100-94}{20}\right)} \\ &= 0.002\ \text{V} \end{aligned}$$

Many current microphone preamplifiers have input impedances that are at least an order of magnitude or larger than the output impedances of commonly encountered microphones. In such instances, Eq. 17-2 can be employed to determine the maximum voltage that a given microphone and sound field will supply to the preamplifier input. The voltage sensitivity of Eq. 17-1 is the one currently employed by most microphone manufacturers.

Another useful sensitivity rating for a microphone is that of power sensitivity. In this instance the focus is placed upon the maximum power that the microphone can deliver to a successive device such as a microphone preamplifier when the microphone is exposed to a reference sound field. In this instance the reference power is one milliwatt or 0 dBm and the reference sound field pressure is one pascal or 94 dB. This rating is more complicated as it involves the microphone output impedance. All microphones regardless of whether the construction is moving coil, capacitor, ribbon, etc. have intrinsic output impedance that in general is complex and frequency dependent. Strictly speaking, in order for such a device to deliver maximum power, it must work into a load that is matched on a conjugate basis with the reactance of the load being the negative of the reactance of the source and the resistance of the load being equal to the resistance of the source (see Chapter 8 *Interfacing Electrical and Acoustic Systems*, Section 8.1, Alternating Current Circuits).

Suppose then that the real part of the microphone's output impedance is R_o. This being the case, the available input power in watts that the microphone can deliver to the input of a successive device, *AIP*, is given by

$$AIP = \left(\frac{1}{4}\right)(E_o{}^2)\left(\frac{1}{R_o}\right) \tag{17-3}$$

If *AIP* is referenced to one milliwatt and the microphone is exposed to a sound field of one pascal then,

$$\frac{AIP}{0.001} = \left(\frac{1}{4}\right)(10^3)\left(10^{\frac{S_V}{10}}\right)\left(\frac{1}{R_o}\right) \tag{17-4}$$

This can be converted to a power level by taking the logarithm to the base ten of Eq. 17-4 and then multiplying by 10 dBm to yield

$$L_{AIP} = (-6 + 30 + S_V - 10\log R_o)\ \text{dBm} \qquad (17\text{-}5)$$

L_{AIP} expresses the power sensitivity of a microphone in terms of dBm/Pa. If our example microphone has an R_o of 200 Ω along with its voltage sensitivity of –60 then its power sensitivity would be

$$-6 + 30 - 60 - 23 = -59\ \text{dBm/Pa}$$

Another useful way to express the power sensitivity of a microphone would be to reference the available input power to a sound field of 0.00002 Pa. This would produce a result 94 dBm lower than that of Eq. 17-5. If we symbolize this rating by G_{AIP} then,

$$G_{AIP} = (S_V - 10\log R_o - 70)\ \text{dBm} \qquad (17\text{-}6)$$

In this rating system the example microphone would produce –153 dBm at the threshold of hearing. The advantage of this system is that the power level supplied by a given talker's microphone is obtained by simply adding G_{AIP} to the pressure level of the talker's voice at the microphone's position. G_{AIP} as defined here is very similar to the EIA rating for microphones. The EIA rating system differs in that rather than employing the actual output resistance of the microphone, a nominal microphone impedance rating is employed instead.

17.3 Thermal Noise

In 1927, while investigating the noise properties of vacuum tube amplifiers, J. B. Johnson of Bell Telephone Laboratories discovered that the input resistance of the subject amplifiers was the source of a noise signal. This resistor noise was found to be separate and apart from other known noise sources associated with vacuum electron tubes. This resistor noise was found to be related to the temperature of the resistor itself and thus was termed thermal noise even though it is also often called Johnson noise in honor of the discoverer. Johnson communicated his experimental data to Harry Nyquist who was a physicist also employed by Bell Labs. Nyquist was able to derive from theory the exact quantitative description of the thermal noise of the resistor and hence this noise source is sometimes also called Nyquist noise.

Later Onsager showed that the thermal noise of a resistor results from a much larger theorem known as the fluctuation-dissipation theorem that is applicable to a larger class of physical systems. In this theorem it is found that the noise signals are properties of only the dissipative elements of a system and not of the energy storage elements. As a consequence, thermal noise in an electrical system is associated only with the real part of the circuit impedance and not with the imaginary or reactive component. Any passive electrical network consisting of resistors, inductors, and capacitors thus contains voltage sources of a thermal origin associated with each of the resistive elements. Furthermore, any passive electroacoustical or electromechanical system at a uniform absolute temperature T can be described by substituting for these systems the corresponding electrical equivalent circuits also at the same absolute temperature.

The fluctuating thermal voltage in a resistor is brought about by thermally induced variations in conduction electron distribution throughout the body of the resistance. An instantaneous non-uniform charge distribution produces an instantaneous voltage difference between the ends of the resistor. Although this voltage difference has an average value of zero, it instantaneously displays a random polarity and size resulting in a non-vanishing root mean square value.

From the Nyquist theorem, each resistor can be represented as a voltage generator in series with the resistance as depicted in Fig. 17-3A where the root mean square voltage of the generator for ordinary temperatures and frequencies is given by

$$V = \sqrt{4kTRB} \qquad (17\text{-}7)$$

where,

k is Boltzmann's constant = 1.38×10^{-23} J/°K,
T is the absolute temperature on the Kelvin scale,
R is the resistance in Ω,
B is the observation bandwidth in Hz.

Fig. 17-3A depicts a thermal noise source whose open circuit voltage between the indicated terminals is given by Eq. 17-7. The spectrum of the noise signal generated by such a source for all practical purposes is independent of frequency meaning that it is a white noise source. Quantum mechanics dictates a frequency dependence of the form

$$\frac{\dfrac{hf}{kT}}{e^{\frac{hf}{kT}} - 1}$$

where,

h is Planck's constant = 6.62×10^{-34} J-s,
T is the absolute temperature,
k is Boltzmann's constant = 1.38×10^{-23} J/°K,

f is the frequency in Hz.

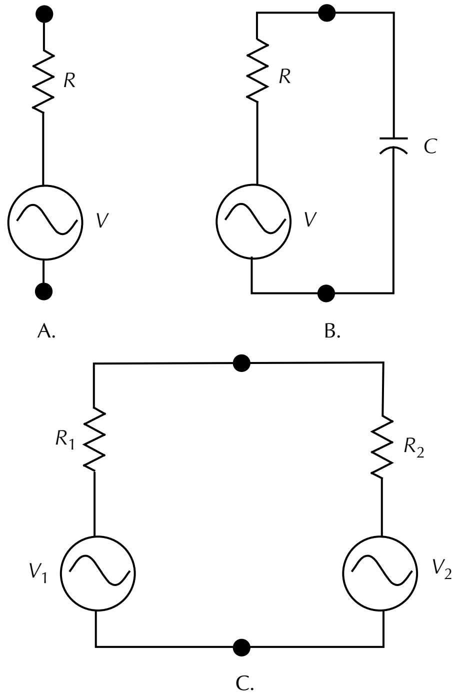


Figure 17-3. Nyquist voltage sources.

This expression begins to differ from unity only at frequencies in the microwave region and beyond and hence can be neglected for the present purposes.

Although thermal noise of a resistor is flat, the frequency dependence is modified by connection to either a capacitor or inductor. In Fig. 17-3B the root mean square noise voltage between the indicated terminals is conditioned by the fact that the RC combination forms a low pass filter. The mean value of the square of the noise voltage generated in the resistor between the frequencies f and $f + df$ can be written as

$$\langle V^2 \rangle = 4kTRdf \tag{17-8}$$

The resistor and capacitor form a voltage divider. The mean square voltage across the capacitor in the frequency interval between f and $f + df$ is obtained by multiplying Eq. 17-8 by the square of the magnitude of the frequency dependent action of this voltage divider. The result of this multiplication is found to be

$$\langle V_C^2 \rangle = \frac{4kTRdf}{4\pi^2 f^2 R^2 C^2 + 1} \tag{17-9}$$

In order to obtain the mean square voltage across the capacitor in a finite frequency interval between the lower frequency f_1 and an upper frequency f_2 it is necessary to integrate Eq. 17-9 over the frequency interval with the result

$$\langle V_C^2 \rangle = \frac{2kT}{\pi C}[\mathrm{atan}(2\pi RCf_2) - \mathrm{atan}(2\pi RCf_1)] \tag{17-10}$$

The root mean square voltage across the capacitor will just be the square root of the expression on the right in Eq. 17-10. This is a general result. When RCf_2 is small the arc tangent terms may be replaced by their arguments with the result that

$$\begin{aligned} \langle V_C^2 \rangle &= 4kTR(f_2 - f_1) \\ &= 4kTRB \end{aligned} \tag{17-11}$$

On the other hand if RCf_2 is quite large with $f_2 \gg f_1$ then the difference in the arc tangent terms is $\pi/2$ and the limiting result becomes

$$\langle V_C^2 \rangle = \frac{kT}{C} \tag{17-12}$$

A more interesting and practical case is that of Fig. 17-3C. Through the application of the principle of superposition, it is possible to determine the circuit current when each source acts alone.

$$\begin{aligned} I_1 &= \frac{V_1}{R_1 + R_2} \\ &= \frac{\sqrt{4kT_1R_1B}}{R_1 + R_2} \\ I_2 &= \frac{V_2}{R_1 + R_2} \\ &= \frac{\sqrt{4kT_2R_2B}}{R_1 + R_2} \end{aligned} \tag{17-13}$$

We are now in position to determine the noise power delivered by each source to the other. Let P_{12} be the noise power delivered by source 1 to source 2 and similarly let P_{21} be the noise power delivered by source 2 to source 1. Then,

$$\begin{aligned} P_{12} &= I_1^{\,2}R_2 \\ &= \frac{4kT_1R_1BR_2}{(R_1+R_2)^2} \\ P_{21} &= I_2^{\,2}R_1 \\ &= \frac{4kT_2R_2BR_1}{(R_1+R_2)^2} \end{aligned} \tag{17-14}$$

It should be observed from Eqs. 17-14 that both lines are the same with the exception of the temperature factor. The source having the lower temperature receives a net power from the higher temperature source regardless of the individual source resistance. Only when temperature equilibrium exists with $T_1 = T_2 = T$ will $P_{12} = P_{21}$. Another point of interest is the noise voltage that appears between the indicated terminals of Fig. 17-3C.

This can also be calculated by superposition. One first calculates the voltage existing there when source 1 acts alolne and then when source 2 acts alone. The voltage when both sources act simultaneously is obtained by the appropriate combination of these two individual voltages. The appropriate combination in this case is the quadratic sum rather than the linear sum as the two generators have completely random phases. The quadratic sum is obtained by taking the square root of the sum of the squares of the individual voltages when acting alone. Let this sum be denoted by V', then

$$V' = \sqrt{\frac{4kBR_1R_2(T_1R_2+T_2R_1)}{(R_1+R_2)^2}} \tag{17-15}$$

If both resistors are at the same temperature, i.e., $T_1 = T_2 = T$ the expression simplifies to

$$V' = \sqrt{\frac{4kBTR_1R_2}{(R_1+R_2)}} \tag{17-16}$$

Clearly, Eq. 17-16 is equivalent to the open circuit voltage of a single resistor whose value is the parallel combination of the two resistors R_1 and R_2. Suppose now that there is a large disparity between the sizes of the two resistors, for instance $R_2 = 100R_1$. In this instance upon dividing R_2 into both the numerator and the denominator underneath the radical Eq. 17-16 becomes

$$\begin{aligned} V' &= \sqrt{\frac{4kTBR_1}{\left(\frac{R_1}{100R_1}+1\right)}} \\ &\approx \sqrt{4kTBR_1} \end{aligned} \tag{17-17}$$

The conclusion for such a situation is that the root mean square noise voltage is essentially that of the significantly smaller resistor in the parallel combination. This is the situation often encountered with present day microphone amplifiers with R_1 corresponding to the microphone's output resistance and R_2 corresponding to the amplifier's input resistance.

Now let the two resistors be equal and each of size R_o. This would be the situation if a microphone were feeding into an amplifier under matched conditions. In this instance

$$V' = \frac{1}{\sqrt{2}}\sqrt{4kTBR_o} \tag{17-18}$$

Eq. 17-18 tells us that the total thermal noise voltage appearing across the amplifier input would be 3 dB less than if the microphone were operating into an open circuit or into a load whose resistance was very much larger than R_o. At the same time, however, the performer's electrical signal generated by the microphone will be reduced by 6 dB. This would indicate that the signal to noise ratio at the system input is worsened by operating into a matched load. This fact should not be considered in isolation however. The important consideration is the signal to noise ratio available from a system consisting of both a microphone and its associated amplifier.

Amplifiers contain several noise sources of their own. In addition to thermal noise associated with circuit resistances there are several noise sources associated with the active amplifying devices themselves. Such noise sources are shot noise, flicker noise, partition noise, and recombination noise. Aside from shot noise whose origin is quantization of electric charge and has a flat spectrum at audio frequencies, the other noise sources do not in general have flat spectra. Even though an amplifier can have several cascaded stages of amplification, in a well-designed system the overall noise behavior is dominated by the first high gain stage of amplification.

A figure of merit with regard to noise behavior of amplifiers can be obtained by considering the degrading effect that the amplifier has upon the signal to noise ratio of the signal source. Fig. 17-4A depicts a signal source consisting of a signal generator of mean square voltage V_S^2 and a source resistance R_S. Along with the source resistance is a thermal noise generator having a mean square voltage in a narrow frequency interval of amount V_n^2.

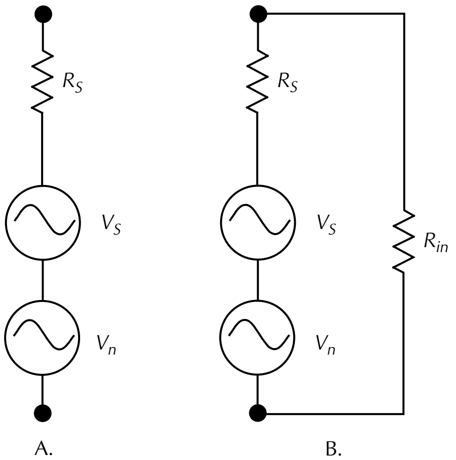


Figure 17-4. Signal source and associated thermal noise.

The ratio of the mean square voltages of the signal and the thermal noise is the best obtainable as allowed by nature and will be called the source signal to noise ratio denoted as SNR_S.The situation changes when the source is attached to the amplifier's input resistance as depicted in Fig. 17-4B. R_{in} affects the operation of the signal source by introducing signal attenuation and contributing thermal noise of its own as discussed previously. Additionally, the presence of R_S affects the operation of the amplifier as it provides a parallel path for amplifier generated input circuit noise currents. This means that the amplifier's noise performance will be sensitive to source resistance.

What matters of course is the signal to noise ratio at the amplifier's output. This ratio can be written as

$$SNR_A = \frac{V_{oS}^2}{V_{on}^2} \tag{17-19}$$

where,

V_{oS}^2 is the amplifier mean square signal output in a narrow band centered on the signal frequency,

V_{on}^2 is the amplifier mean square noise output in a narrow band centered on the signal frequency.

Finally, the figure of merit of noise performance can be written as

$$F = \frac{SNR_S}{SNR_A} \tag{17-20}$$

Most audio amplifiers in current use have input resistances that are considerably larger than the usual source resistances. This being the case,

$$\begin{aligned} SNR_S &= \frac{V_S^2}{4kTR_S\Delta f} \\ SNR_A &= \frac{A^2 V_S^2}{4kTR_S\Delta f + V_{nA}^2} \end{aligned} \tag{17-21}$$

In Eq. 17-21 the term A^2 is the square of the voltage amplification of the amplifier and V_{nA}^2 is the mean square noise voltage at the amplifier output contributed by the amplifier alone when operating from a noiseless source resistance equal to R_S. The expression noiseless source resistance requires some explanation.

Recall from the discussion leading to Eq. 17-17 that the thermal noise at the amplifier output in this case is fully accounted for by the first term in the denominator of the amplifier's signal to noise ratio expression. Additionally, the size of the second term hinges in part on the size of the source resistance independent of the resistor's thermal noise. It is convenient to refer the amplifier's output noise contribution to its input circuit by picking a suitably sized mean square noise voltage, V_{nin}^2, such that

$$V_{nA}^2 = A^2 V_{nin}^2 \tag{17-22}$$

Finally, we will pick a resistor whose thermal noise mean square voltage is equivalent to that of the amplifier when referred to the amplifier's input.

$$4kTR_e\Delta f = V_{nin}^2 \tag{17-23}$$

The noise figure of merit now becomes

$$\begin{aligned} F &= \frac{A^2 4kTR_s\Delta f + A^2 4kTR_e\Delta f}{A^2 4kTR_S\Delta f} \\ &= \frac{R_S + R_e}{R_S} \end{aligned} \tag{17-24}$$

A word of caution is called for here. Unlike thermal noise that has a flat spectrum, amplifier associated noise is dependent on frequency so R_e is itself frequency dependent.

Remember also that the size of R_e is different for different choices for the source resistance. The noise figure of an amplifier is denoted by NF and is related to the figure of merit discussed here.

$$NF = 10\ \text{dB}\log(F) \tag{17-25}$$

In the foregoing discussion we considered a narrow band of frequency centered about the frequency of the signal source. The noise figure as well as the equivalent noise resistor so derived is that for the operating frequency of the signal source. Notice, however, that the expression for the noise figure does not contain any factor involving the size of the genuine signal.

In the experimental determination of noise figure one simply requires that a resistance equal to the intended source resistance be connected directly to the amplifier input. This resistor brings along its associated thermal noise as required. This resistor's temperature must be maintained fixed at whatever absolute temperature is required. The total noise signal at the output of the amplifier must pass through a narrow bandpass filter centered on the particular frequency of interest before being observed on a true rms voltmeter of sufficient sensitivity. The voltage gain of the amplifier must be accurately known or measured and the noise contribution of the source resistor can be calculated.

Special measurement techniques may be required because Δf is quite small and the amplifier noise observed in a narrow frequency interval is consequently also small. If one is interested in the average noise figure over the entire audio band, then one employs a maximally flat bandpass filter of appropriate width. Regardless of whether the observation bandwidth is small, Δf, or broad, B, the appropriate width is not the frequency interval between the half power points as conventionally defined for the following reason.

The thermal noise spectrum is flat and the frequency interval employed in those equations is for an ideal filter. Depending on the steepness of the slopes of a conventional filter significant noise power may be passed for frequencies outside of the conventional pass band. The width of such filters must be decreased until they pass the same power as would be passed by an ideal filter in the frequency interval of interest when supplied with a flat spectrum.

The noise properties of amplifiers may be reported in a related but different fashion from that of *NF.* This is done by reporting the amplifier's equivalent input noise expressed as a power level relative to one milliwatt. This expression is called the *EIN*. The *EIN* in dBm is calculated from

$$EIN = NF + 10\log(kTB) + 30 \tag{17-26}$$

EIN in effect is the available input power level of a fictitious noise generator having an open circuit voltage equal to the quadratic sum of that produced by the source resistance and the equivalent noise resistance of the amplifier. The internal resistance of this fictitious noise generator is set equal to the source resistance. If you subtract the noise figure of the amplifier expressed in dB from *EIN* you arrive at the thermal noise available input power level of the source resistance alone.

If interest is centered on a system's noise floor, then *EIN* is the answer without further calculation. If interest is directed toward degradation of signal to noise ratio (*SNR*), the important quantity to know is *NF.* The signal to noise ratio expressed in decibels at the amplifier output is less than that of the signal to noise ratio of the signal source by an amount equal to *NF.* Upon employing the notation of Eq. 17-21

$$10\ \text{dB}\log(SNR_A) = 10\ \text{dB}\log(SNR_S) - NF \tag{17-27}$$

At this point it is worthwhile to examine some sample calculations. One must start with both the microphone specifications and the amplifier specifications appropriate to the source impedance presented by the microphone as well as the pressure level of the talker.

Talker	Microphone	Amplifier
$L_P = 80\,\text{dB}$	$S_V = -60\,\text{dB}$	$R_S = 200\,\Omega$
	$G_{AIP} = -153\,\text{dBm}$	$NF = 5\,\text{dB}$
	$R_o = 200\,\Omega$	$B = 20{,}000 - 20 = 19{,}980\,\text{Hz}$

The object is to determine the signal to noise ratio at the amplifier output expressed in dB. The calculation will be carried out by employing two different techniques. First it should be observed that the output resistance of the microphone serves as the source resistance for the amplifier and that the amplifier data is appropriate for use with this microphone.

First, employ Eq. 17-2 to determine the talker's root mean square signal voltage produced by the microphone.

$$\begin{aligned} E_o &= 10^{\left(\frac{-60+80-94}{20}\right)} \\ &= 10^{\left(\frac{-74}{20}\right)} \\ &= 199.5\ \mu\text{V} \end{aligned}$$

Next employ Eq. 17-7 to determine the root mean square thermal noise voltage produced by the microphone's output resistance.

$$\begin{aligned} V &= \sqrt{(4)(1.38)(10^{-23})(293)(200)(19{,}980)} \\ &= 0.25\ \mu\mathrm{V} \end{aligned}$$

In the above step the ambient temperature was taken as 20°C which corresponds to 293° on the absolute or Kelvin temperature scale. The *SNR* of the signal source expressed in decibels is then

$$\begin{aligned} 20\ \mathrm{dB}\log\left(\frac{E_o}{V}\right) &= 20\ \mathrm{dB}\log\left(\frac{199.5\times10^{-6}}{0.254\times10^{-6}}\right) \\ &= 57.9\ \mathrm{dB} \end{aligned}$$

Finally, the signal to noise ratio at the amplifier output expressed in dB from Eq. 17-27 is

$$\begin{aligned} 10\ \mathrm{dB}(SNR_A) &= 57.9\ \mathrm{dB} - 5\ \mathrm{dB} \\ &= 52.9\ \mathrm{dB} \end{aligned}$$

Before we begin the alternative calculation it should be pointed out that when Eq. 17-4 is applied to a thermal noise source and then referenced to one milliwatt the result expressed as an available noise power level is given by

$$L_{AIPNoise} = [10\log(kTB) + 30]\ \mathrm{dBm} \qquad (17\text{-}28)$$

The noise available input power level of the microphone is then

$$\begin{aligned} &10\log[(1.38)(10^{-23})(293)(19{,}980)] + 30 \\ &= -130.9\ \mathrm{dBm} \end{aligned}$$

The microphone produces a signal available input power level that is the sum of G_{AIP} and L_P. This results in a signal whose available input power level is

$$-153 + 80 = -73\ \mathrm{dBm}$$

The source signal to noise ratio in decibels is just the difference between these two levels or

$$-73 - (-130.9) = 57.9\ \mathrm{dB}$$

The signal to noise ratio at the amplifier output expressed in dB is then

$$\begin{aligned} 57.9\ \mathrm{dB} - NF &= 57.9\ \mathrm{dB} - 5\ \mathrm{dB} \\ &= 52.9\ \mathrm{dB} \end{aligned} \qquad (17\text{-}29)$$

This mediocre result is a consequence of an insensitive microphone combined with a noisy amplifier.

The noise figure of an amplifier based on bipolar junction transistors is strongly dependent on the resistance of the signal source. Such amplifiers perform at their best with source resistances falling in the range of $5\,\mathrm{k}\Omega$ to $20\,\mathrm{k}\Omega$. Any given amplifier has an optimum source resistance for which its noise figure is the least. It appears then that optimum noise performance will be obtained with low impedance microphones when their output resistance can be made to appear to be the optimum source resistance required by the amplifier at hand.

Noise figures as low as 1 dB are obtainable in this fashion. In such cases it is reasonable to employ a quality step up transformer of the appropriate turns ratio at the amplifier input. A step up transformer with a turns ratio of 1:n will transform R_o by a factor of n^2. The signal voltage from the microphone as well as the thermal noise voltage of the microphone are increased by the same factor n. The signal to noise ratio of the microphone is thus unchanged provided that the amplifier's input resistance as viewed from the microphone is still large compared with R_o.

The foregoing analysis is somewhat simplified in that no consideration was given to the possible existence of reactances being associated with the microphone's output impedance. This is reasonably accurate with regard to capacitor microphones that incorporate low noise field effect transistor source followers internal to the microphone housing. This internal circuitry operates from either local battery or externally applied phantom power. These source followers offer output impedances of the order of $50\,\Omega$ that is purely real and hence resistive. The microphone cable does introduce shunt capacitance and this capacitance in conjunction with R_o forms a first order low pass filter. For reasonable cable lengths, the corner frequency of this filter is well beyond the audio spectrum and does not impact signal to noise calculations.

Further treatment is required when dynamic microphones are involved. Consider a typical dynamic microphone. The capsule of the microphone is similar to a miniature loudspeaker. The capsule has a diaphragm and an attached coil that moves in the air gap of a magnetic structure. The diaphragm's motion is driven by acoustic pressure variations impinging on its surface. The coil typically has a resistance of only a few ohms and an inductance that is a small fraction of a millihenry. A small step up transformer is included in the microphone housing that transforms the resistance to a few hundred ohms and the inductance to a few millihenries. The capsule has a low frequency mechanical resonance that influences the impedance behavior of the overall assembly. With the exception of scale, the output impedance magnitude has a dependence on frequency similar to a dynamic loudspeaker as depicted in Fig. 18-22 of Chapter 18 *Loudspeakers and Loudspeaker Arrays.*

The manufacturer's rated impedance is usually the magnitude found at the minimum of the curve located above the mechanical resonance frequency. For frequencies well above the mechanical resonance, the microphone's output impedance is that of a series resistance R_o and a series inductance L_o shunted by the capacitance C_c of any attached microphone cable. This structure is depicted in Fig. 17-5A. This structure as viewed from the vantage point of the microphone amplifier is displayed in Fig. 17-5B.

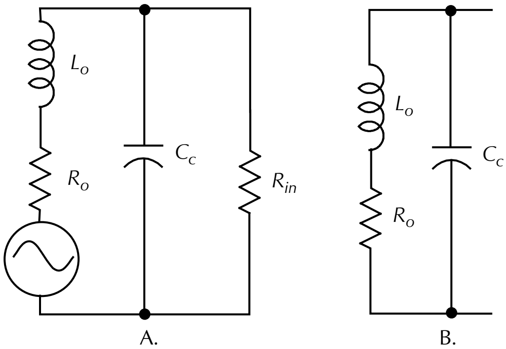


Figure 17-5. Dynamic microphone equivalent circuit and amplifier source impedance.

As depicted in Fig. 17-5A the internal impedance of the dynamic microphone along with the shunt capacitance of the cable connecting the microphone to the amplifier form a second order low pass filter that is terminated by the amplifier's input resistance. The open circuit voltage of the microphone is represented by the voltage generator of size V_o. Typical values for the circuit components are $R_o = 200\,\Omega$, $L_o = 2\,\text{mH}$, $C_c = 2\,\text{nF}$, and $R_{in} = 1\,\text{M}\Omega$. The amplitude behavior of this circuit is best described by plotting the attenuation versus frequency where the attenuation is $20\text{dB}\log(V_{in}/V_o)$ and V_{in} is the voltage appearing across the amplifier input resistance. This result appears as Fig. 17-6.

The filter's behavior is innocuous up to about 2 kHz. Above 2 kHz the filter begins to exhibit a gain rather than a loss. This is the result of an under damped resonance that occurs at a frequency far above the normal audio band. The filter gain at 20 kHz is less than 0.6 dB and is actually slightly beneficial in compensating for the natural high frequency fall off of the microphone's natural response. The microphone's talker signal as well as its thermal noise signal experiences this same gain and signal to noise behavior is not affected. The other consideration with regards to noise behavior is that of the source impedance as viewed from the standpoint of the amplifier. This impedance is depicted in Fig. 17-5B. The magnitude of this source impedance is plotted in Fig. 17-7.

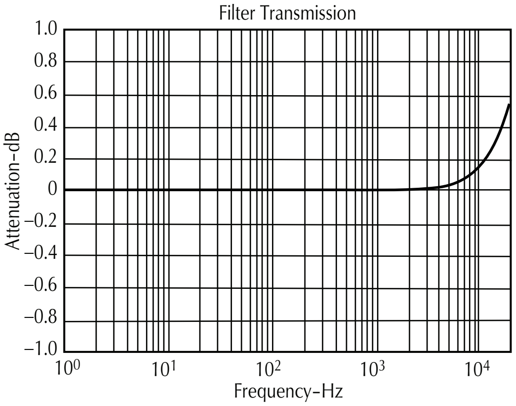


Figure 17-6. Filter amplitude behavior.

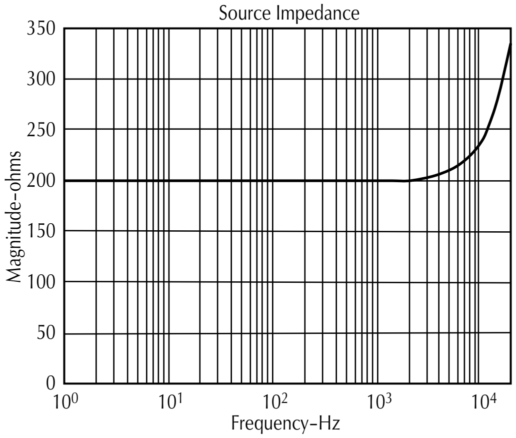


Figure 17-7. Source impedance viewed by amplifier.

The source impedance is resistive at a constant value of 200 Ω until about 2 kHz. Above that point up to 20 kHz the source impedance has a frequency dependent reactive component in addition to the resistance of 200 Ω. This behavior will have a modest effect on the signal to noise ratio in the frequency decade from 2 kHz to 20 kHz. This modest effect may well be an improvement as most amplifiers prefer source impedances greater than 200 Ω. In summary, neglecting the reactive component of the microphone's output impedance apparently leads to only minor errors in signal to noise ratio calculations.

17.4 Microphone Selection

Microphones are usually selected on the basis of mechanism, sensitivity, nature of response, polar response pattern, and handling characteristics. Mechanism refers to the physical nature of the transducing element of the microphone. Sensitivity in current practice refers to the voltage sensitivity S_V. Nature of response refers to whether the microphone output is proportional to acoustic pressure, acoustic pressure gradient, or acoustic particle velocity. Polar response patterns summarize a microphone's directional characteristics. Handling characteristics are a result of whether or not the structure of the microphone housing is mechanically isolated from the transducing structure of the microphone. Below is a listing of popular microphones according to the transducing mechanism.

1. Carbon.
2. Capacitor.
3. Moving coil.
4. Ribbon.
5. Piezoelectric.

Carbon

Carbon microphones made their advent as transmitters in early telephones. Pressure variations on a metallic diaphragm actuated a metallic button contact to either increase or decrease the compaction of carbon granules contained in a brass cup so as to decrease or increase the resistance of the assembly. The impinging sound thus modulated the direct current in a circuit containing a battery and the microphone element. Carbon microphones are quite sensitive and inexpensive to construct. In addition to the normal thermal noise, such microphones suffer from fluctuations in contact resistance between carbon granules even in the absence of acoustic excitation. The high noise floor and restricted frequency response limit the application of such microphones in sound reinforcement systems.

Capacitor

Capacitor microphones exist in two basic forms. In one form a capacitor has a front plate formed by a flexible low-mass, metallic, or metal film diaphragm separated by an air gap from an insulated, rigid metallic perforated back plate. Air motion through the perforations in the back plate serves to damp the mechanical resonance of the diaphragm. This resonance occurs at a high frequency as a result of a stiff, low mass diaphragm. The diaphragm is operated at ground potential while the back plate is charged through a very high resistance by a dc voltage source ranging up to 200 V.

In a second form, a permanently polarized dielectric or electret is positioned on the surface of the back plate removing the necessity for an external polarizing voltage source. In both instances the capacitor circuit is completed through a resistance of the order of $10^9\,\Omega$ and the charge on the capacitor remains approximately constant. Pressure variations on the flexible diaphragm produce changes in the air gap dimension thus raising or lowering the capacitance by a small amount depending upon the degree of diaphragm displacement. With a constant charge on the variable capacitor, the voltage variations track the diaphragm displacement variations.

The capacitor circuitry itself is of high impedance and requires that a field effect transistor (FET) source follower be contained within the microphone housing. The source follower may be energized by a local battery in the case of the electret form or may derive its power from the polarizing voltage source in the pure air capacitor form. These microphones, though not the most rugged, can be of extremely high quality with regard to frequency response. As will be discussed later, the construction details of the microphone capsule may be varied to make the microphone capsule sensitive to either acoustic pressure or acoustic pressure gradient.

Moving Coil

The moving coil microphone and the ribbon microphone are collectively referred to as being dynamic microphones. Much discussion has been given previously with regard to some of the features of the moving coil microphone. The mechanical resonance of the moving coil structure is usually made to occur at the geometric mean of the low frequency and high frequency limits describing the microphone's pass band. In a typical case this resonance occurs at about 630 Hz. In the pressure responsive version of such a microphone the back chamber to the rear of the diaphragm contains an acoustic resistance that highly damps the diaphragm mechanical resonance. This damping greatly broadens the resonance forcing the response to be uniform except at the frequency extremes.

Oftentimes a small resonant tube tuned to a low frequency and vented to the outside is incorporated in the rear cavity. In addition to extending the response at low frequencies, this tube allows the static air pressure in the rear chamber to track slow changes in atmospheric pressure. Even in microphone structures featuring an otherwise sealed rear cavity a slow leak must always be provided for static pressure equalization. A small air chamber that is resonant at a high frequency may also be located in the rear cavity in order to enhance response at high frequencies. Moving coil microphone structures are usually quite rugged.

Ribbon

The ribbon microphone employs a conductor in a magnetic field as does a moving coil microphone. Unlike the moving coil, which is located in a radially directed magnetic field, the conductor in a ribbon microphone is a narrow, corrugated metal ribbon located in a linearly directed magnetic field that is perpendicular to the length of the ribbon. The ribbon itself constitutes the diaphragm, both faces of which are exposed to external sound fields.

The driving force on the ribbon is directly proportional to the pressure difference acting on the two faces of the ribbon and hence is proportional to the space rate of change of acoustic pressure. The space rate of change of pressure is called the pressure gradient. The ribbon responds to the acoustic particle velocity with maximum response occurring when the incident sound is normal to a face of the ribbon. This microphone is inherently directional with a figure eight polar pattern. Though featuring excellent performance over a wide frequency range, the structure is inherently fragile and is not suitable for exterior use under windy conditions.

Piezoelectric

Piezoelectric microphones depend on a structural property possessed by certain dielectric crystals and especially prepared ceramics. The nature of this property is that if the crystal or ceramic is subjected to a mechanical stress, its shape will be distorted. When this occurs, an electric field appears in the substance as a result of shifted ion positions within the structure. A capacitor can be formed employing such a dielectric that will generate a voltage that is proportional to the mechanical stress. The mechanical stress can be made to result from the motion of a diaphragm exposed to acoustic pressure. In this fashion it is possible to construct a relatively simple, inexpensive pressure sensitive microphone. Piezoelectric microphones have very high capacitive output impedances. In the past the high voltage sensitivity of such microphones made them popular for recorders and simple public address applications where quite short connecting cables were possible. They are still employed in some sound level meters but other professional application is quite restricted.

Matching Talker to Microphone

Distant or bashful talkers require microphones of higher voltage sensitivity in order to produce voltage levels matching those required by microphone input amplifiers. Nearby and professional talkers require microphones of less sensitivity in order to match amplifier input requirements without the use of pads in the input circuitry. Rock singers are an extreme case requiring the least input sensitivity and further requiring both breath blast and pop filters particularly when pressure gradient microphones are employed. Table 17-1 lists representative voltage sensitivity ranges typical of microphones classified according to mechanism.

Table 17-1. Microphone Sensitivity Comparison

Microphone Mechanism	S_V in dBV/Pa Range
Carbon	–20 to 0
Capacitor	–50 to –25
Dynamic	–60 to –50
Piezoelectric	–40 to –20

17.5 Nature of Response and Directional Characteristics

Pressure microphones are those where only one side of the diaphragm is exposed to the actuating sound field. Such devices are basically insensitive to the direction of the arriving sound as long as the wavelength is large compared with the diaphragm circumference. At high frequencies when the wavelength becomes comparable to or even less than the diaphragm circumference, two directional effects become evident. For sound directly incident on the exposed face of the diaphragm, the partial reflection of the pressure waveform at the diaphragm surface increases the acoustic pressure amplitude over that which would exist in an undisturbed sound field. For sound incident from the rear of the exposed face of the diaphragm, the active face of the diaphragm is in the shadow of the microphone’s housing structure and experiences a pressure less than that of the undisturbed sound field. This front to back discrimination can only be avoided by employing physically small microphone structures. This is the reason why measurement microphones often have capsules of ¼ inch diameter or even less.

A controlled directional response can be obtained by employing a sensing diaphragm both faces of which are exposed to the sound field of interest. Such diaphragms experience a driving force that depends on the spatial rate of change of pressure rather than on the pressure itself. Consider the situation shown in Fig. 17-8.

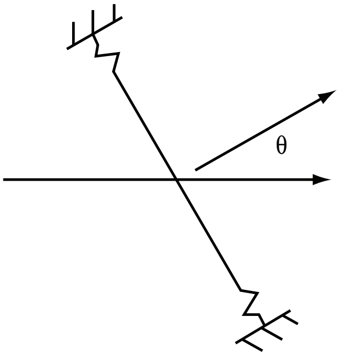


Figure 17-8. Compliantly mounted diaphragm with both sides exposed to sound field.

Fig. 17-8 is a bare bones illustration of a diaphragm stripped of details of the transducing mechanism. Both sides of the diaphragm are exposed to a sound wave that is propagating along the horizontal axis. The diaphragm may be circular as in a capacitor or moving coil microphone or rectangular as in a ribbon microphone. The principal axis of the microphone is directed perpendicular to the plane containing the diaphragm and as illustrated forms an angle θ with the direction of the incident sound. When θ has the value $\pi/2$, both faces of the diaphragm experience identical pressures and the net driving force on the diaphragm is zero. Now when θ is 0, the sound wave is incident normally on the diaphragm and the driving force on the left face of the diaphragm will be the pressure in the sound wave at the left face's location multiplied by the area of the left face.

The diaphragm material however is not porous so sound must follow an extended path around the diaphragm along which the sound pressure can undergo a change before reaching the right face. The net driving force on the diaphragm will be the difference in the pressures on the two faces multiplied by the common diaphragm surface area. The pressure difference can be calculated by taking the product of the space rate of change of acoustic pressure, known as the pressure gradient, with the effective acoustical distance separating the two sides of diaphragm. The least value of this distance is the diaphragm diameter in the case of a circular diaphragm.

For a ribbon diaphragm the appropriate value would approximate the geometric mean of the diaphragm's length and width. Details of a particular microphone housing structure that provide a baffle- like mounting will tend to increase the effective separation. If θ is not zero, the microphone axis is inclined to the direction of the incident sound and the pressure difference is lowered according to the cosine of the angle. Before reading the following mathematical analysis, it might be useful to review the material on phasor mathematics in Chapter 6 *Mathematics for Audio Systems*.

As a first case consider that the sound source is quite distant from the microphone location so that that the incident sound can be described by a plane wave. The mathematical description of such a wave where the direction of propagation is that of the x-axis is

$$p(x,t) = p_m e^{j(\omega t - kx)} \tag{17-30}$$

where,

p_m = the acoustic pressure amplitude,
ω = angular frequency = $2\pi f$,
k = propagation constant = $\omega/c = 2\pi/\lambda$,
c = phase velocity,
λ = wavelength.

Under this circumstance, the net driving force acting on the diaphragm in the direction of increasing x is given by evaluating the following expressions with x set equal to the coordinate of the diaphragm's center.

$$\begin{aligned} F(t) &= \left[p(x,t) - \left\{p(x,t) + \frac{\partial}{\partial_x}p(x,t)d\cos\theta\right\}\right]S \\ &= \left[-\frac{\partial}{\partial_x}p(x,t)d\cos\theta\right]S \end{aligned} \tag{17-31}$$

where,

S is the surface area of one side of the diaphragm,
$(\partial/\partial_x)p(x,t)$ is the gradient of the acoustic pressure in the direction of increasing x.

The pressure gradient is calculated by taking the partial derivative with respect to x of Eq. 17-30 as follows

$$\begin{aligned} \frac{\partial}{\partial_x}p(x,t) &= -jkp_m e^{(\omega t - kx)} \\ &= -jkp(x,t) \end{aligned} \tag{17-32}$$

Upon substituting the result of Eq. 17-32 into Eq. 17-31, the driving force becomes

$$F(t) = jkp(x, t)Sd\cos\theta \tag{17-33}$$

In a given sound wave of normally encountered intensities there exists a relationship between the acoustic pressure and the acoustic particle velocity. The ratio of the acoustic pressure to the particle velocity is called the specific acoustic impedance of air for the wave type in question. This ratio for plane waves is a real number equal to the normal density of air multiplied by the phase velocity of sound. One then can substitute for the acoustic pressure in Eq. 17-33 in terms of the particle velocity to obtain an alternative expression for the driving force.

$$\begin{aligned} F(t) &= jk\rho_0 cu(x, t)Sd\cos\theta \\ &= j\omega\rho_0 u(x, t)Sd\cos\theta \end{aligned} \tag{17-34}$$

The significance of the imaginary operator j in this equation simply means that the phase angle of the driving force leads that of the particle velocity by $\pi/2$ radians or 90°. The amplitude of the driving force would be

$$F_m = \omega\rho_0 u_m Sd\cos\theta \tag{17-35}$$

where,
u_m is the particle velocity amplitude.

The more often encountered case is where the source is nearby to the microphone location. In such an instance the appropriate wave description is that of a spherical wave propagating along a radial line from the sound source. Mathematically, such a wave is described by

$$p(r, t) = \frac{A}{r}e^{j(\omega t - kr)} \tag{17-36}$$

where,
$\frac{A}{r} = p_m$ is the pressure amplitude that is now position dependent,
A is a constant determined by the sound source.

The pressure gradient is now more complicated as the space variable r appears in both the denominator and the exponent of the expression for the acoustic pressure.

$$\frac{\partial}{\partial_r}p(r, t) = -\left(\frac{1}{r} + jk\right)p(r, t) \tag{17-37}$$

If the center of the diaphragm is located at a distance r from the sound source then the driving force on the diaphragm for the spherical wave becomes

$$F(t) = \left(\frac{1}{r} + jk\right)p(r, t)Sd\cos\theta \tag{17-38}$$

The driving force now has two components, one of which is in phase with the acoustic pressure while the other leads the acoustic pressure by 90°. The specific acoustic impedance of air for spherical waves is not as simple as was the plane wave case. The ratio of the acoustic pressure to the particle velocity is now

$$\frac{p(r, t)}{u(r, t)} = \frac{j\omega\rho_0}{\frac{1}{r} + jk} \tag{17-39}$$

Upon solving Eq. 17-39 for the acoustic pressure in terms of the particle velocity and substituting into Eq. 17-38 one obtains the very important result

$$F(t) = j\omega\rho_0 u(r, t)Sd\cos\theta \tag{17-40}$$

The importance of this result is apparent when Eq. 17-40 is compared with Eq. 17-34. With the exception of the identity of the space variable, the two equations are identical implying that pressure gradient microphones respond to the particle velocity in exactly the same fashion whether the incident sound wave is plane, spherical, or a combination of the two. In contrast, pressure sensitive microphones respond to acoustic pressure whether the source is nearby (spherical case) or distant (plane case). In fact, for a pressure sensitive microphone the driving force depends only on the acoustic pressure and is given by the direction independent expression

$$F(t) = p(r, t)S \tag{17-41}$$

Another very important aspect of pressure gradient microphones is the proximity effect. This phenomenon becomes apparent by a rearrangement of Eq. 17-39. This equation is solved for the particle velocity in terms of the pressure and the terms then multiplied in both numerator and denominator by the radial distance while making use of the fact that $k = \omega/c$ to obtain

$$u(r, t) = \frac{1 + jkr}{jkr}\left(\frac{p(r, t)}{\rho_0 c}\right) \tag{17-42}$$

The significance of this result is more pronounced when one examines the magnitude of the particle velocity.

$$|u(r, t)| = \frac{\sqrt{1 + k^2 r^2}}{kr}\left(\frac{|p(r, t)|}{\rho_0 c}\right) \qquad (17\text{-}43)$$

When the radial distance is large or the wavelength is short or of course both of these are true, then Eq. 17-43 reduces to $|u(r, t)| = |p(r, t)|/\rho_0 c$ with the significance that the particle velocity is directly proportional to the acoustic pressure. On the other hand, when r is small or the wavelength is large or a combination is true, the reduction becomes $|u(r, t)| = (1/\omega r)(|p(r, t)|/\rho_0)$ with the significance that the particle velocity varies inversely with frequency. As a consequence, when a sound source is in close proximity to a pressure gradient microphone the lower frequencies of the source produce a larger response than do the higher frequencies. This is the basis for the proximity effect.

One final observation with regards to the directional characteristics of pressure gradient microphones. From Eq. 17-38, when θ is in the range $\pi/2 < \theta < 3\pi/2$, the cosine of θ is itself a negative quantity and the polarity of the driving force as well as the electrical output signal of the microphone is reversed. A use will now be made of this fact in discussing a microphone structure that possesses a variety of several different directional patterns.

A structure consisting of both a pressure gradient microphone element and a pressure microphone element makes possible a microphone possessing adjustable directional characteristics. The elements should individually be small and located close together with the diaphragms of the two elements located in the same plane. A single signal based on a linear sum of the signals from the individual elements is generated by the combination. The root mean square open circuit electrical output of the assembly can be written as

$$E_o = \alpha(\beta + \gamma\cos\theta) \qquad (17\text{-}44)$$

where,

α is a dimensional constant,

β is the fraction of the pressure microphone electrical signal,

γ is the fraction of the pressure gradient microphone electrical signal,

θ is the angle of incidence of the acoustic signal.

The fractional signals can be formed and summed through the employment of passive circuitry contained within the microphone housing. The polar response curve of the microphone for a given choice of coefficients is obtained by allowing θ to range continuously from 0 to 2π while plotting the curve

$$r = |\beta + \gamma\cos\theta| \qquad (17\text{-}45)$$

where,

r is the radial distance from the origin and has a maximum value of 1,

β and γ are fractional coefficients with $\beta + \gamma = 1$,

θ is the angle of incident sound relative to principal axis of microphone.

Although β and γ are arbitrary within the constraint that they sum to unity, there are particular values that have proven to be quite useful. This information is listed in Table 17-2.

Table 17-2. Polar Pattern Parameters for Microphone Directional Characteristics

Polar Pattern	β	γ	*RE**	*DF**
Omni	1	0	1	1.0
Gradient	0	1	0.33	1.7
Subcardioid	0.7	0.3	0.55	1.3
Cardioid	0.5	0.5	0.33	1.7
Supercardioid	0.37	0.63	0.268	1.9
Hypercardioid	0.25	0.75	0.25	2.0

*Based on data from Shure Incorporated.

Some practitioners prefer to employ directional microphones because such microphones respond to reverberant acoustical power arriving from all directions with reduced sensitivity as compared with the same acoustical power arriving along the principal axis of the microphone. This property is expressed by the entry labeled *RE* in Table 17-2. *RE* stands for random efficiency. The hypercardioid pattern, for example, has a random efficiency of ¼. The response to power distributed uniformly over all possible directions is thus only ¼ that for the same total power arriving on axis.

The entry labeled *DF* in Table 17-2 compares the working distance of a directional microphone to that of an omnidirectional microphone. The *DF* for a hypercardioid microphone is 2 meaning that the working distance for a source on axis for this microphone can be twice as large as that for an omni in order to achieve the same direct to reverberant sound ratio in the output signal.

These factors when considered alone would lead one to believe that higher gain before acoustic feedback instability would be achievable through the employment of directional microphones. This is not necessarily the case. As a class, omnidirectional microphones exhibit smoother frequency responses than do directional microphones. The frequencies of oscillation triggered by acoustic feedback, the ring frequencies, depend upon a number of factors.

Prominent causative agents are peaks in microphone response and peaks in loudspeaker response coupled with antinodes in the normal modes of the room. Room modes at even moderate frequencies can be quite dense. As a consequence, a single peak in either microphone or loudspeaker response may trigger an entire chorus of slightly different ring frequencies. This set of facts would tend to favor omnidirectional microphones over directional ones. The deciding factor is usually not immunity to feedback from the reverberant field but rather the necessity to reject a nearby source of objectionable sound including possible strong discrete reflections.

A microphone consisting of a separate pressure and pressure gradient element is quite versatile in that it offers all of the polar response patterns listed in Table 17-2 assuming that it contains the appropriate switch selectable passive circuitry necessary to properly combine the signals from the individual elements. Such a microphone, however, inherently has a shortcoming in that the centers of the two elements are physically offset.

Sound waves incident on the device in other than the principal plane arrive at the two elements at slightly different times. The difference in arrival times introduces a phase difference between the electrical signals generated by the two elements. This phase difference can be significant at high frequencies and can distort the directional response pattern in the high frequency region. Fortunately, it is possible to avoid the offset problem through the design of a single diaphragm device that also has useful directional characteristics.

Fig. 17-9 is a bare bones illustration of a compliantly mounted diaphragm and a back enclosure that is vented through a porous screen to the external environment. The diaphragm may be part of either a capacitor or moving coil type of transducer, the details of which are not shown for simplicity. A sound wave is incident on the left face of the diaphragm. The direction of the incident wave makes an angle θ with the principal axis of the system. The principal axis is perpendicular to the plane that contains the diaphragm. The acoustic pressure on the left face of the diaphragm assuming a spherical wave is given by

$$p_1 = \frac{A}{r}e^{j(\omega t - kr)} \tag{17-46}$$

The center of the porous screen to the right of the diaphragm is separated from the corresponding point at the center of the diaphragm by an acoustical distance that amounts to $(d + L)$ where d is the diameter of the diaphragm. We need now to calculate the acoustic pressure at a point just to the right of the center of the porous screen. The acoustic pressure in the incident wave on the diaphragm is a known quantity p_1. As was done in the case of the pressure gradient microphone, we first calculate the rate of pressure change with distance along the direction of propagation. Next we find the component of this change in the direction of interest. Finally we multiply this component by the acoustical distance between the points of interest. This last step yields the pressure change. What is desired of course is the pressure at the second point. This is the pressure at the initial point plus the change in pressure. Upon letting p_2 represent the acoustical pressure at a point immediately to the right of the center of the porous screen then,

$$p_2 = p_1 + (d + L)\cos\theta\frac{\partial}{\partial_r}p_1 \tag{17-47}$$

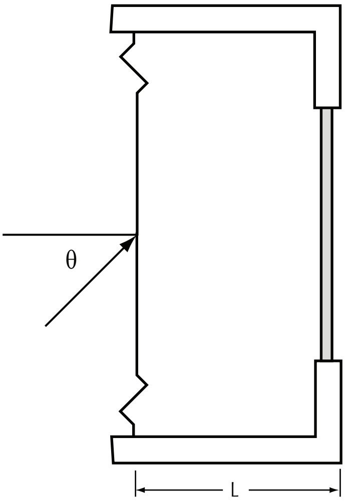


Figure 17-9. Simplified illustration of a single diaphragm that is sensitive to a combination of pressure and pressure gradient.

The driving force that actuates the diaphragm, however, is the pressure difference between p_1 and the pressure in the cavity to the rear of the diaphragm multiplied by the surface area of one side of the diaphragm. A detailed analysis would show that the pressure in the cavity, p_e, depends upon both p_1 and p_2. Recall that for a pressure sensitive microphone the diaphragm driving force is directly proportional to the acoustic pressure whereas for a pressure gradient microphone it is directly proportional to the gradient of the acoustic pressure.

In the capsule described above the driving force on the diaphragm is proportional to a linear combination of the pressure and pressure gradient terms. The sizes of the coefficients in the linear combination and consequently the particular directional polar pattern hinge on the volume of the cavity, the areas occupied by the diaphragm and the porous screen, the mechanical properties of the diaphragm, and the porosity of the screen. Such microphones are usually constructed having a dedicated directional pattern. The majority of the cardioid family of directional microphones is constructed in this fashion.

Most microphones have cylindrical symmetry and basically circular diaphragms. The principal axis of such a microphone is centered on the diaphragm, perpendicular to the plane of the diaphragm, and directed along the cylindrical axis as illustrated in Fig. 17-10. The directional polar pattern in a plane is obtained by varying the angle of incident sound relative to the principal axis of the microphone.

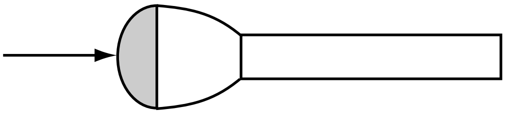


Figure 17-10. Illustration of the principal axis of a cyindrically symmetric microphone.

The three dimensional directional response of such a microphone is obtained by revolving the directional polar pattern about the cylindrical axis of the microphone. Ribbon microphones, however, don't follow the above rules, as their diaphragms do not possess cylindrical symmetry. Such microphones are usually designed to be addressed from the side as illustrated in Fig. 17-11.

The directional response in the horizontal plane of the depicted ribbon microphone is a figure eight. Revolving this pattern about the principal axis generates two spheres that describe the microphone's response in three dimensions. The polar directional patterns listed in Table 17-2 are displayed in Fig. 17-12A while the three dimensional directional response is sketched in Fig. 17-12B.

The polar patterns of Fig. 17-12A are the theoretical ideals and have a linear radial axis consistent with the form of the describing equations. Real microphones fall short of the theoretical ideal in two ways. They never display complete nulls in response and the polar response curves are frequency dependent. Compare the measured polar response curves of a cardioid microphone presented in Fig. 17-13 with its counterpart in Fig. 17-12A.

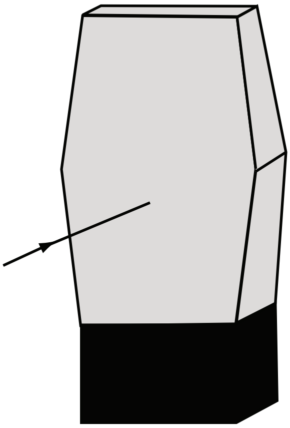


Figure 17-11. Position of the principal axis of a classic ribbon microphone.

Manufacturer's polar response data is usually presented employing a logarithmic polar axis while excluding a small region in the vicinity of the origin. Such a presentation for yet again a different cardioid microphone is given in Fig. 17-14.

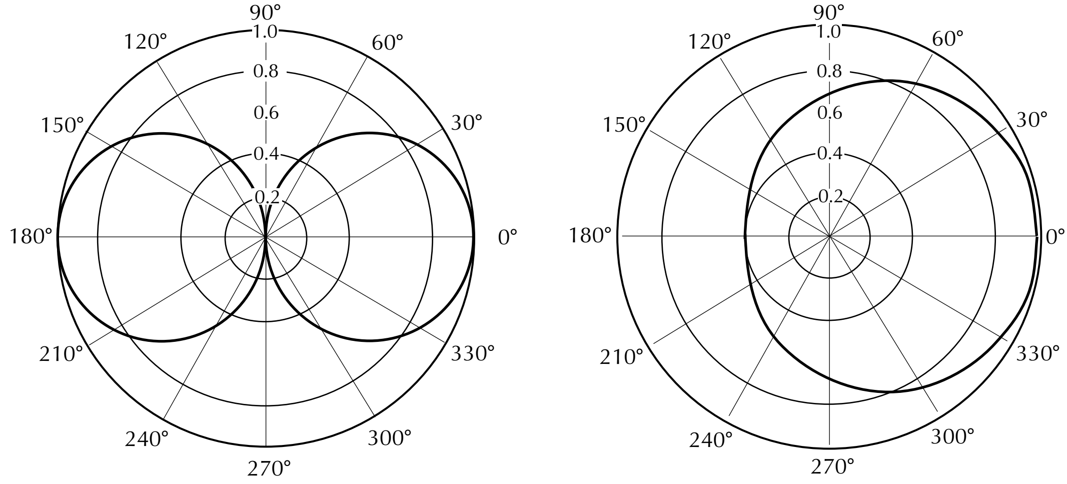
In examining Fig. 17-14 note that the reference axis has a different orientation and that the radial coordinate represents attenuation expressed in decibels relative to the on axis value.

17.6 Boundary Microphones

All microphones can suffer anomalies in their responses depending upon the relative location of the sound source and nearby reflecting surfaces. A commonly encountered incidence of this is illustrated in Fig. 17-15.

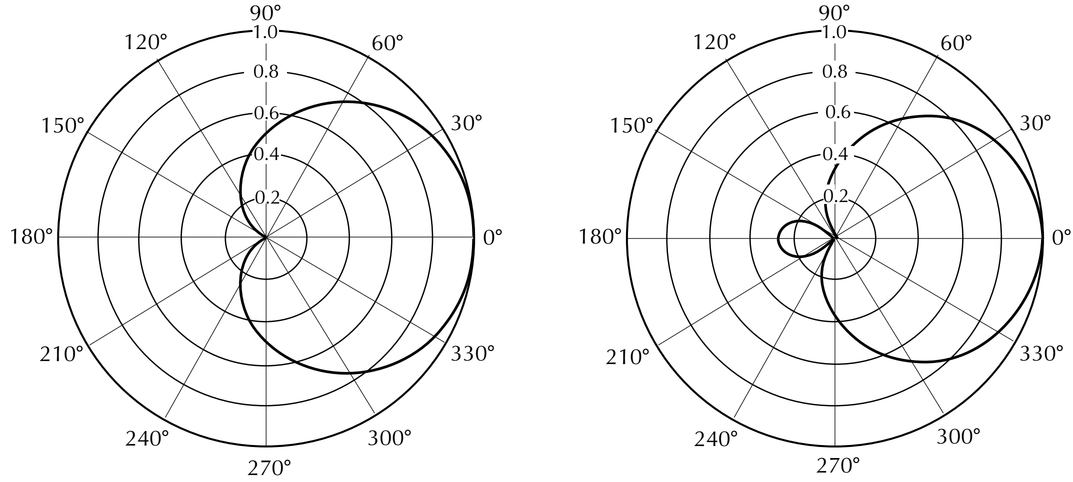
The combination of the direct sound with the late arriving and slightly attenuated reflected sound produces a comb filter amplitude response pattern. The notches in the pattern occur at those frequencies where the two signals differ in phase by an odd integral multiple of π radians. This corresponds to the total distance traveled by the reflected sound being an odd integral multiple of $\lambda/2$ greater than that traveled by the direct sound. Such a condition results in destructive interference. If d_r is the total distance traveled by the reflected sound in arriving at the microphone and d_d is the distance traveled by the direct sound in reaching the microphone, then the notches occur at those wavelengths where

$$\lambda_n = \frac{2}{n}(d_r - d_d) \tag{17-48}$$



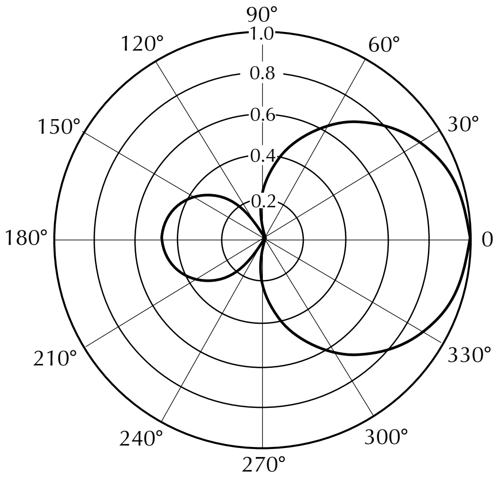
Pressure Gradient Pattern.

Subcardioid Polar Pattern.



Cardioid Polar Pattern.

Supercardioid Polar Pattern.



Hypercardioid Polar Pattern.

Figure 17-12A. Standard polar patterns.

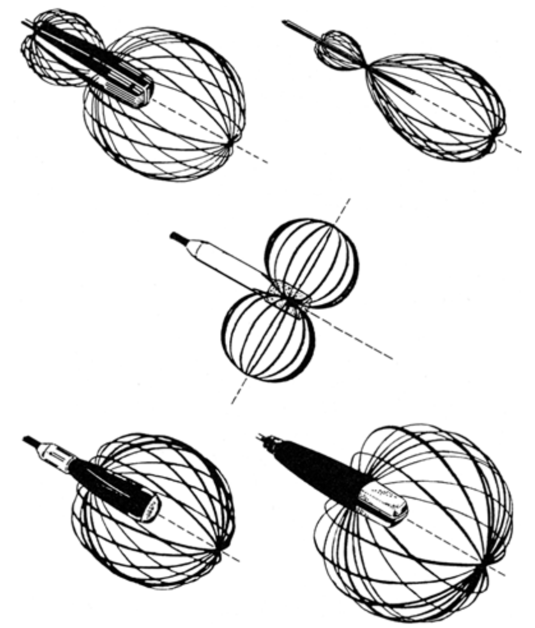


Figure 17-12B. Microphone three dimensional directional response. (*Courtesy Shure Incorporated.*)

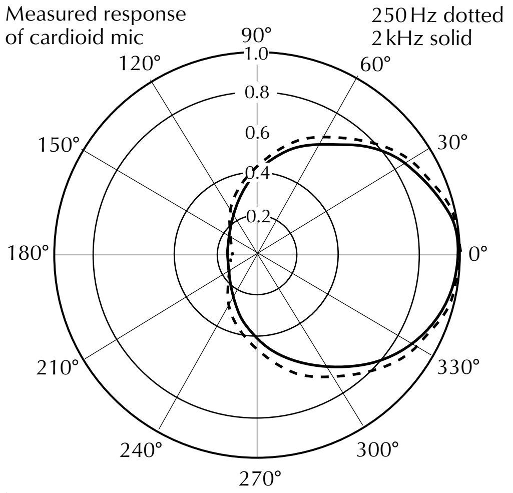


Figure 17-13. Measured polar response of cardioid microphone at 250 Hz and 2 kHz.

where,

n is any odd integer.

The interference frequencies corresponding to these wavelengths can be easily calculated by making use of the fact that $f = c/\lambda$. This yields

$$f_n = n\frac{c}{2(d_r - d_d)} \tag{17-49}$$

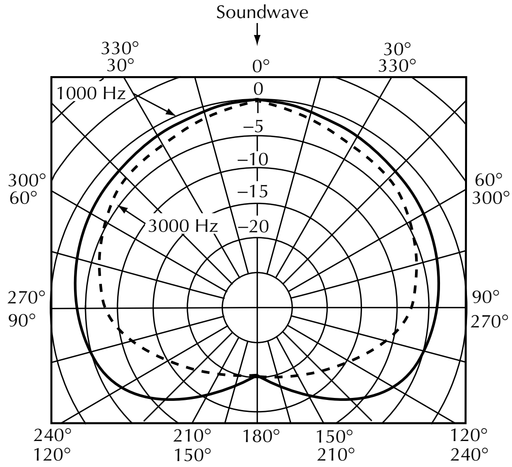
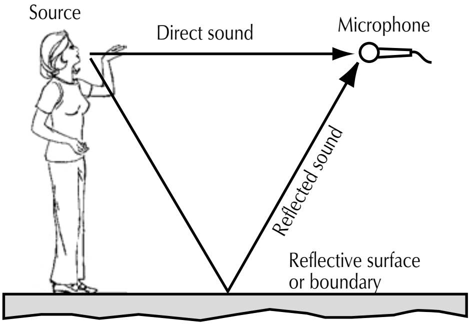
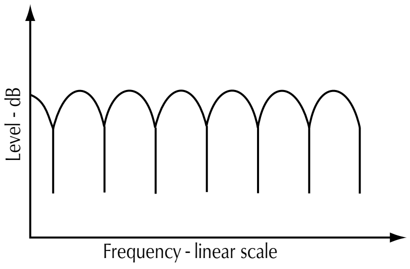


Figure 17-14. Polar response of a cardioid microphone.



A. Situation



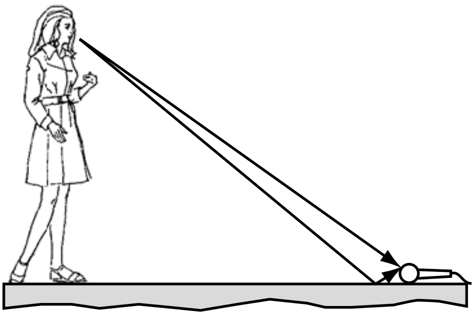
B. Frequency response

Figure 17-15. Microphone that receives direct sound and delayed reflected sound. (*Courtesy Bruce Bartlett.*)

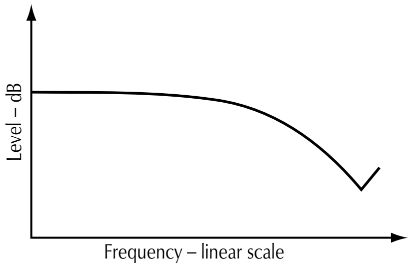
The lowest interference frequency occurs for $n = 1$. The next interference frequency has $n = 3$. These frequencies are

$$\begin{aligned} f_1 &= \frac{c}{2(d_r - d_d)} \\ f_3 &= 3\frac{c}{2(d_r - d_n)} \end{aligned} \tag{17-50}$$

Three conclusions are worthy of note. The difference between adjacent interference frequencies is a constant equal to $2[c/2(d_r - d_d)]$, this separation is just twice the value of the lowest interference frequency, and the smaller the path difference between the reflected and direct sound, the higher the frequency of the first interference notch. The fact that the difference between successive interference frequencies is a constant means that this phenomenon is best viewed on an analyzer with a linear frequency axis. Additionally, reducing the path difference between the reflected and direct sound shifts the first interference frequency to a higher value. This is illustrated in Fig. 17-16.



A. Situation



B. Frequency response

Figure 17-16. Microphone on floor that receives direct sound and slightly delayed reflected sound. (*Courtesy Bruce Bartlett.*)

Referring to Fig. 17-15 upon taking the direct and reflected paths to form an equilateral triangle and assuming a typical height for a young girl, the difference between the direct and reflected path lengths will be a little over 5 ft. The first interference notch will occur at about 100 Hz. The notch separation would then be about 200 Hz. Contrast this with the circumstance of Fig. 17-16 where the path length difference could be as small as an inch. The first notch would now occur between 6 kHz and 7 kHz. This suggests that it might be possible to make the first interference notch occur at a frequency above 20 kHz thus placing it beyond the audio band.

Suppose that a pressure sensitive microphone is oriented to face the reflecting surface directly while having the diaphragm only 0.25 inch from the boundary. The path length difference is now also of the order of 0.25 inch. The first interference notch now falls between 25 kHz and 30 kHz. Such considerations as well as others led Edward Long and Ronald Wickersham to point out the need for a pressure calibrated capsule when used in the pressure zone near a boundary.

Kenneth Wahrenbrock built the first commercially successful microphone systems to incorporate the basic principles enunciated by Long and Wickersham. Ken employed a miniature electret capacitor microphone element with the capsule's diaphragm being spaced a millimeter or less above a mounting plate. This small spacing places the first interference notch at more than two octaves above the audio spectrum thus allowing a smooth amplitude response throughout the normal audio band. Microphones constructed in this manner are called pressure zone microphones or PZMs. Fig. 17-17 illustrates the structure of a PZM where the intent is to mount the entire plate on a larger plane boundary.

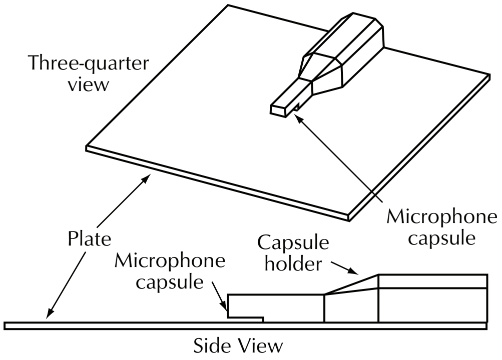


Figure 17-17. Construction of a typical PZM. (*Courtesy Bruce Bartlett.*)

At this juncture it is appropriate to consider the conditions imposed on acoustic wave propagation at a boundary surface. Initially we will consider that the boundary is perfectly rigid. With this in mind, examine the situation depicted in Fig. 17-18.

When the spherical source is radiating at a single frequency, the air particles are oscillating back and forth along lines directed radially outward from the surface of the source while the pressure alternates a small amount above and below normal atmospheric pressure. This disturbance away from static conditions begins at the surface of the small spherical source and propagates outward along radial lines. The wavefronts are spherical in shape and represent surfaces all points of which have the same phase.

The drawing of Fig. 17-18 depicts in two dimensions an instantaneous snapshot of this situation where the spherical wavefronts appear in cross

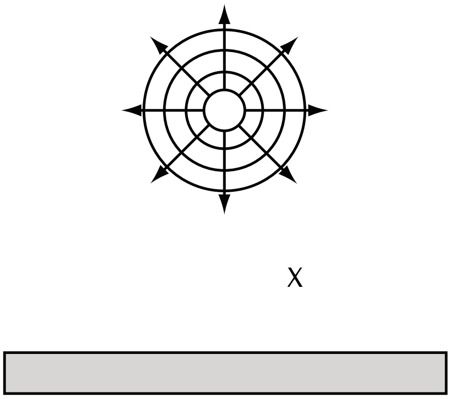


Figure 17-18. Simple spherical source in the presence of a rigid boundary.

section as concentric circles with the radial spacing between successive circles being arbitrarily set equal to the wavelength indicating a phase change of 2π from one circle to the next. As time goes on each spherical wavefront will expand so that its radius will increase by one wavelength as the elapsed time increases by one period. The air particle velocity is a vector oscillating along any radial line while the acoustic pressure has no direction and is a scalar quantity.

We assume a simple spherical source for the purpose of analysis because even complicated acoustical fields can be described by a superposition of distributed simple spherical sources of adjustable strengths, positions, and phases. The chore of calculating the acoustic pressure at an arbitrary point such as X in the drawing is quite difficult in the general case where the boundary may not be rigid and may have arbitrary dimensions.

The problem is greatly simplified when the boundary is considered to be perfectly rigid, its dimensions are large compared with the wavelength, and the source is not positioned near the boundary edge. When the boundary is perfectly rigid there can be no air particle motion at the boundary surface that is perpendicular to the boundary surface. At the boundary surface then, the normal component of the air particle velocity must be zero implying that this component of the air particle velocity must be reflected with a reversal of polarity or a phase change of π.

A plane wave, for instance, could have complete normal incidence at the boundary and would be reflected back on itself while undergoing a reversal in its direction of propagation. The reversal of the particle velocity that occurs at the rigid boundary fully accounts for the reversal in the direction of propagation provided that the pressure is reflected without a change in phase. When the conditions of simplification are satisfied, the problem can readily be analyzed by the method of images as depicted in Fig. 17-19.

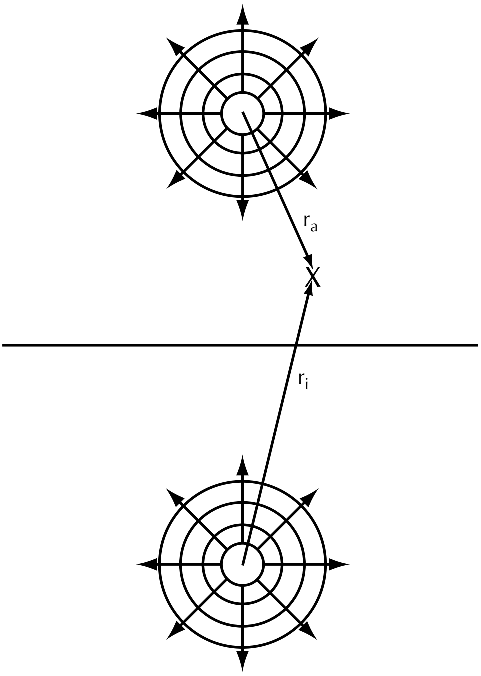


Figure 17-19. Boundary solution by the method of images.

In the method of images one replaces the rigid physical boundary by an image plane at the boundary location and an image source positioned just as far below the plane as the actual source is above it. The properties of the image source are taken to be identical to those of the actual source. The acoustic pressure for any point in the half space on or above the image plane is readily calculated by adding the contribution of the actual source to that of the image source. The acoustic pressure at the point X for example is given by

$$\begin{aligned} p &= p_a + p_i \\ &= \left(\frac{A}{r_a} e^{j(\omega t - kr_a)} + \frac{A}{r_i} e^{j(\omega t - kr_i)}\right) \end{aligned} \tag{17-51}$$

The composite signal of Eq. 17-51 is exactly the same as the direct wave to point X from the actual source combined with the wave which would have been reflected from the physical boundary. This composite signal will display all of the usual comb filter effects. When the point X is located on the image plane, the distances r_a and r_i become equal indicating that the composite pressure becomes twice that of the actual source operating in a free field at the same distance. This doubling of pressure at the image plane surface is the consequence of the pressure in the reflected wave being in phase with the pressure of the incident wave. Such a situation is depicted in Fig. 17-20.

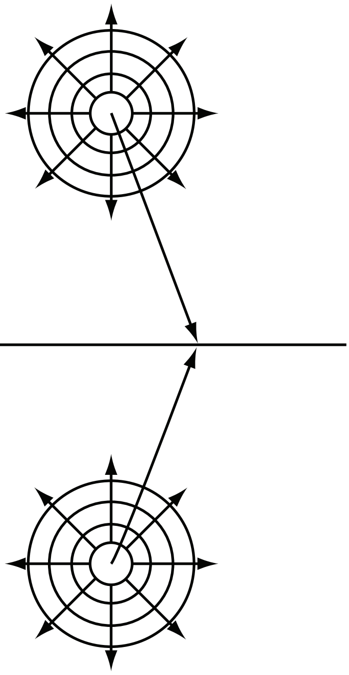


Figure 17-20. Observation point located on the image plane.

When the observation point is located on the image plane, the normal or vertical components of the particle velocity from the two sources are equal in magnitude but oppositely directed forcing the net normal component to be zero thus satisfying the required boundary condition for a rigid boundary. On the other hand, the horizontal or tangential components of the particle velocity contributed by the two sources are equal in magnitude and in the same direction. The net tangential particle velocity is thus twice as large as it would have been for a single source operating in a free field at the same radial distance.

A solution to a problem in wave acoustics involves finding a mathematical function that satisfies three conditions. First and foremost, the function must satisfy the wave equation. Secondly, it must satisfy the conditions that exist at the surface of the source or sources. Thirdly, it must satisfy the conditions that exist at any and all boundaries.

The boundary condition at the surface of a rigid boundary is that the normal component of particle velocity must be zero. The method of images is constructed so that it forces satisfaction of the condition on the normal component of particle velocity. Fortunately, there exists a mathematical uniqueness theorem that is applicable to problems in wave acoustics. This means that once one finds the function that fulfills the three conditions listed above it will be the sole solution to the problem.

The exact doubling of the acoustic pressure and tangential component of the particle velocity only occurs on the surface of a truly large, rigid boundary. Real surfaces of course are not perfectly rigid and may also be porous to some degree in which case the boundary conditions will be different and the pressure increase will be less than ideal. In such instances a detailed knowledge of the specific acoustic impedance of the boundary surface is necessary in predicting the exact behavior.

Furthermore, we assumed that the boundary dimensions were large compared with the source wavelength. Even with a rigid boundary, a 6 dB increase in pressure above free field conditions will occur only for those frequencies where the linear dimensions are large compared with the wavelength.

When the dimensions of the boundary are comparable to the wavelength the pressure increase begins to fall and becomes equal to the free field value when the wavelength is large compared with the dimensions of the boundary. Shelving filters may be employed to compensate for this behavior.

17.7 Wireless Microphones

Modern wireless microphones allowing untethered motion of the user have proven themselves to be indispensable in concerts, religious services, dramatic arts, and motion picture or video production.

Wireless microphones for use in the performing arts and sound reinforcement first made their appearance about 1960. The first transmitter units were designed to operate in the broadcast FM band between 88 MHz and 108 MHz. The receivers were conventional FM broadcast units. The transmitters did not have to be licensed as the low radiated powers involved complied with Part 15 of the FCC rules. Frequency modulation was accomplished in the transmitter by allowing the audio voltage signal to vary the junction capacitance of a bipolar transistor connected as a Hartley or other simple oscillator tuned to the desired carrier frequency in the FM band. Such oscillators were prone to drift in operating frequency as the transistor characteristics were sensitive to both temperature and supply voltage variations. This required periodic retuning of the receiver to compensate for transmitter frequency drift. This was particularly true of the very early units that employed germanium transistors. Significant improvement in this regard was made possible with the availability of suitable silicon transistors. One of the authors well remembers hand crafting several body pack transmitters in 1965 for use by lecturers at Georgia Tech. The receivers employed were H. H. Scott units that had been modified to incorporate automatic frequency control circuitry to compensate for the transmitter drift within reasonable limits. These early units had acceptable audio bandwidths but the simple modulation technique employed did not produce large frequency deviations resulting in a small dynamic range of the recovered audio signal.

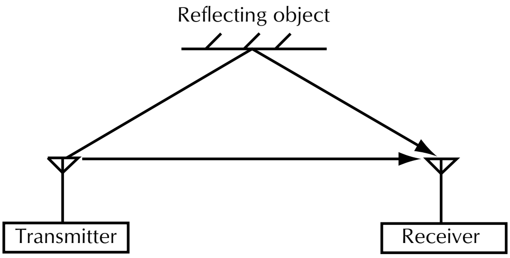
Those of us who have experienced the entire history of wireless microphones consider the present day versions to be truly remarkable. Not only have the early shortcomings been addressed but also features not even envisioned by the early practitioners have been added. The modern history of wireless microphone technology can be divided into two periods. The first period corresponds to the time when commercial television broadcasting employed only analog techniques and the second period commenced with the required changeover to digital television broadcasting techniques. In the first period frequency space was made available in both the VHF and UHF frequency bands and the wireless microphones all employed analog technology. In the second period of wireless microphone employment frequency allocation is restricted to UHF and the most recent wireless microphone transmitters and receivers employ digital technology. At the beginning of the second period wireless microphone manufacturers had to surmount many technical problems. Fortunately, they did their jobs very well and managed to produce digital transmitters and receivers with increased dynamic range as well as stability and ease of management as compared with the former analog systems. Additionally, required receiving antenna lengths are much more manageable in the UHF band. For example, with a carrier frequency of 900 MHz and a wave speed of 3×10^8 m/s, the wavelength becomes one third of a meter or about 13 inches. The required receiving antennas range between ¼ to ½ wavelength and thus have lengths falling between about 3 to 6 inches. Both analog and digital wireless systems have many features in common and both systems remain in use where allowed. This discussion will begin with the analog systems of the first period of employment.

There are several significant technical innovations incorporated in current analog wireless microphone systems that are worthy of note. Each of these will be discussed in turn.

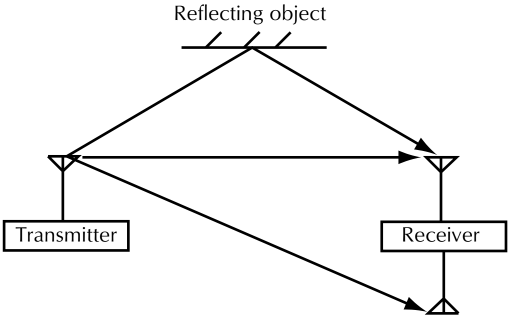
1. Receiver Assisted Setup.
2. Space Diversity Reception.
3. Transmitter Pre-emphasis–Receiver De-emphasis.
4. Transmitter Compression–Receiver Expansion.

A difficult problem associated with setting up wireless microphone systems in the past has been that associated with determining interference free operating frequencies. This was particularly true when the application required the simultaneous operation of a large number of separate audio channels each of which required an individual radio frequency assignment. Receivers having assisted setup facilities have built in protocols for scanning the entire operating band and identifying those potential operating frequencies that are free of any radio frequency carrier at the time of scan. Several such scans performed over a period of time usually are quite successful in defining interference free operating frequencies.

Space diversity reception solves a problem depicted in Fig. 17-21A by means of an arrangement suggested by Fig. 17-21B.



A. Multipath Reception



B. Space Diversity Reception

Figure 17-21. Multipath and space diversity reception.

In Fig. 17-21A a single receiving antenna is employed. This antenna receives a signal via a direct path to the transmitter as well as a transmitter signal that has been reflected by a nearby object and thus follows a longer more indirect path along its way to the receiving antenna. The phases of these two signals having the same frequency are different and hence they can interfere with each other. The interference may be either constructive or destructive according to the degree of phase difference. When the interference is destructive, the resultant signal may be so weak that the receiver will not be able to recover the program material. The arrangement shown in Fig. 17-21B greatly reduces the probability that there will be a complete loss of program material. In this arrangement two antennas located somewhat less than a wavelength apart are employed. In this arrangement the reflected signal may not even arrive at the second antenna as shown. Even when this is not the case or when there are other reflecting objects, the chances that both antennas are subjected to destructive interference simultaneously is greatly reduced.

There are several techniques for handling the signals that appear in the space diversity antennas. In one technique the space diversity receiver is fitted with separate radio frequency amplifiers for each antenna. The signals from each of these amplifiers is compared as to strength with the stronger signal at any instant being switched to the remainder of the single receiver circuitry. In a variation on this technique, the signals from both radio frequency amplifiers are summed and then fed to the rest of the circuitry of a single receiver with no switching being involved. Lastly, two receivers set to receive the same carrier frequency are employed, one for each receiving antenna. The automatic gain control voltages that are developed at each receiver's detection stage are compared with the audio output circuitry being switched to that of the receiver having the larger control voltage. This last technique is the most expensive and even though it involves switching has perhaps the best performance overall.

Analog wireless microphone transmitters employ a relatively small frequency deviation in the frequency modulation process. The modulation index is thus small. This restricts the dynamic range that is available for program material and weak signals may be lost in the noise floor. A long term average of the spectral density associated with both voice and music programs exhibit a broad maximum in the vicinity of 500 Hz accompanied by a roll off in density beyond about 2 kHz. The spectral density is the average power per unit frequency interval. This being the case, it is necessary to pre-emphasize the higher frequencies in the audio material prior to further signal processing. The normal range of the audio material to be transmitted may well be as large as 80 dB while the available range in the small deviation FM transmitter may be only 40 dB. The 80 dB range of the audio material is squeezed into the 40 dB range available by 2 into 1 compression prior to the modulation process. After transmission and reception at the receiver, the recovered audio material occupying a 40 dB range is first subjected to a 1 into 2 expansion in order to restore the full dynamic range of 80 dB. This is then followed by a de-emphasis of the audio material above 2 kHz in order to restore the natural spectral balance of the audio material. Fig. 17-22 displays typical pre-emphasis and complimentary de-emphasis curves with the upper curve being that of pre-emphasis. The combination of the two yields flat response across the audio band.

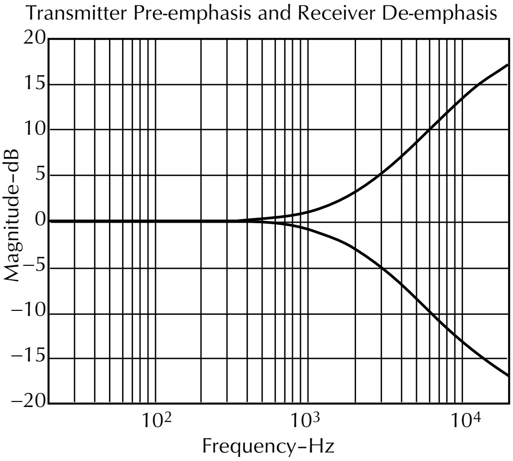


Figure 17-22. Typical pre-emphasis and de-emphasis curves.

The process of compressing the audio dynamic range prior to transmission and expanding the range of the audio material following reception has been termed compansion. A typical compression curve employed in the audio circuitry of the transmitter followed by the complimentary expansion curve employed in the audio circuitry of the receiver are displayed in Fig. 17-23.

Transmitter units may be hand-held with a built in microphone element or a body pack unit provided with a mini receptacle for a microphone connection. The microphones employed with body pack units are usually miniature dynamic or electret capacitor microphones attached to short cables fitted with mating connectors to that of the transmitter. The microphone elements are fitted with clips for attachment to the user's clothing. Occasionally, the microphone element may be part of a head microphone boom structure.

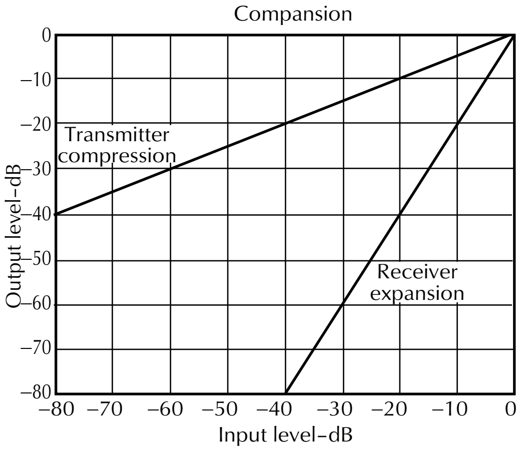


Figure 17-23. Dynamic range compression at the transmitter and complimentary expansion at the receiver.

Typical transmitter features are:

1. Power on-off switch.
2. Carrier frequency selection and indicator.
3. Battery level indicator.
4. Audio gain control.
5. Audio overload indicator.
6. Audio mute switch on body pack units.
7. 9 volt battery.

Receiver units may be stand-alone or rack-mounted and are usually powered from conventional power mains. Audio outputs are provided at both line and microphone level.

A typical space diversity receiver providing assisted setup has the following features:

1. Power on-off switch.
2. Scan or operate control.
3. Carrier frequency indicator.
4. Squelch control.
5. Active receive antenna indicator.
6. Radio frequency level indicator.
7. Transmitted audio level indicator.
8. Transmitter battery life indicator.
9. Audio output level control.

One final note with regard to analog wireless microphone systems that has been distilled from years of sad personal experience. The first three rules for dealing with wireless microphone systems are:

1. Batteries.
2. Batteries.
3. Batteries!

Wireless transmitters are usually powered by 9 volt batteries that may be composed from primary or non-rechargeable cells or secondary cells that are rechargeable. Even if one ordinarily uses rechargeable batteries, it is well to keep a fresh supply of non-rechargeable units on hand. The histories of rechargeable batteries must be carefully managed in order to assure their proper performance. Many practitioners prefer to employ only fresh non-rechargeable batteries along with frequent replacement because of sad experiences with rechargeable units. Battery failure at a critical moment can lead to years of bad dreams.

Although details vary dependent on the various manufacturers a high quality digital wireless system would typically feature 24 bit digital audio along with AES standard encryption for privacy, ethernet networking setup capability for multiple receivers, and spectral scanning capability for choosing interference-free operating frequencies. Some advanced systems even offer frequency diversity operation for remotely changing the operating frequency of body packs or handheld transmitters while in use in the case of a transiently interrupting signal. Both hand held and body pack transmitters offer selectable radiated power in the range of one to twenty mW. The outstanding feature of the digital systems is the increased available dynamic range that can have an A weighted value in excess of 120 dB. The mobile transmitter units usually employ rechargeable Li-ion batteries. A photograph of a space diversity digital wireless microphone system complete with both hand held and body pack transmitters is presented in Fig. 17-24.

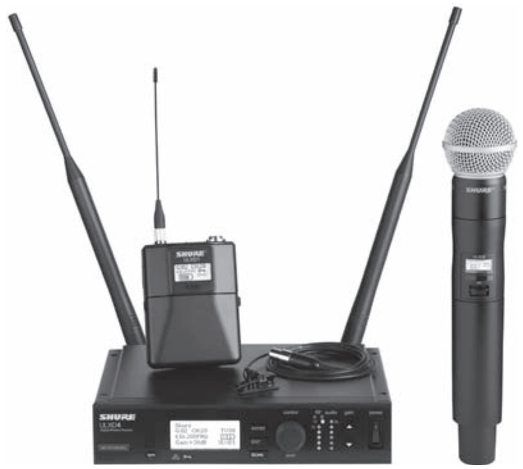
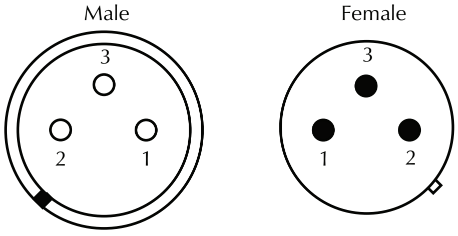


Figure 17-24. Wireless microphone system. (*Courtesy Shure Incorporated.*)

17.8 Microphone Connectors, Cables, and Phantom Power

It is almost universal practice in professional audio to provide signal sources with male connectors and signal receivers with female connectors. Additionally, it is common practice to employ balanced circuits for both input and output in those instances where the signal levels are low and susceptible to electrical noise or cross talk interference. Indeed, many systems maintain balanced linking circuits throughout regardless of the signal levels.

If one excludes miniature microphones that constitute a special case, the de facto standard microphone connector is the XLR-3. The male and female versions of this connector are illustrated in Fig. 17-25.



XLR-3 Connectors

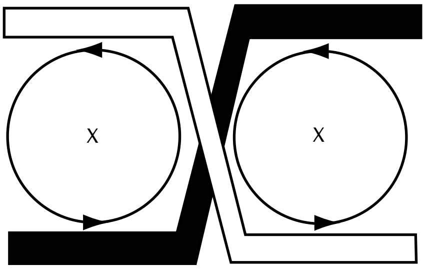
Figure 17-25. Pin arrangements of XLR-3 connectors.

Through the years the assignment of functions to the various pins has varied. The present standard assignment of the male connector at the microphone has pin 1 connected to the microphone case. Pin 2 is connected to the microphone circuitry such that a positive pressure on the microphone diaphragm drives the voltage at pin 2 in the positive sense. Pin 3 is connected to the microphone circuitry such that a positive pressure on the microphone diaphragm drives the voltage at pin 3 in the negative sense. Pins 2 and 3 are balanced with respect to pin 1.

Quality microphone cables consist of a twisted pair of insulated, color-coded inner conductors formed from stranded copper wire covered by a tightly woven copper braided shield with the combination encased in an insulating jacket. The conductors may be tinned although this is not always the case. The cable is fitted with a female connector at one end and a male connector at the other. The connector pin assignments in this instance have pin 1 connected to the shield with the option of also strapping the connector shell to pin 1. Pin 2 is connected to the positive signal conductor while pin 3 is connected to the negative signal conductor. Microphone cable is also often used as the connecting cable in link circuits between mixers, subsequent signal processing units, and power amplifiers.

Shielded, twisted pairs in balanced circuits are an absolute necessity in handling low level signals in order to avoid electromagnetic interference. The braided shield alone offers protection from electrostatic fields but offers very little protection from changing magnetic fields. The practice of employing twisted pair conductors stems from experience gleaned from the early days of the telephone industry.

In former times long distance circuits between cities and local circuits in rural areas employed open bare wire pairs affixed to separate glass insulators attached to multiple cross arms which were in turn elevated by poles. It was learned early on that open-air electrical power lines that often followed parallel paths caused interference. It was found that by periodically transposing the positions occupied by the two conductors of a given circuit pair that the interference could be greatly reduced if not eliminated altogether. This transposition amounted to periodically twisting without touching one conductor of a circuit pair over the other, in effect forming an insulated, twisted pair even though the distance between twists was relatively large. The explanation for this annulment of the interference appears in Fig. 17-26.



Twisted pair

Figure 17-26. Twisted pair exposed to a time changing magnetic field.

In Fig. 17-26 imagine that the twisted pair of conductors is replicated both to the right as well as to the left to form an extended circuit. Imagine also that in the vicinity a magnetic field is instantaneously directed into the figure as indicated by the X’s and that the field strength is increasing with time. Examine the two closed paths as indicated by the circles. According to Lenz’s law, the induced voltage acting in the loops has the sense indicated by the arrows. Now look at the white conductor in the upper left, the induced voltage in this portion of conductor is in the same direction as is the arrow adjacent to it. Compare that with the induced voltage in the white conductor in the lower right in which the induced voltage is oppositely directed.

The same analysis applied to the two similar segments of the black conductor yields identical results. There is no voltage induced in either conductor in the transposition region as the arrows in the adjacent circles are oppositely directed. In practice, the magnetic field alternates but as it changes its direction of growth, the induction in the loops reverses direction also while the net voltage induced in the transposed conductors remains at zero. Static magnetic fields are of no consequence unless a conductor is moving through them. Even so, a twisted pair translated through a magnetic field that is static in time will experience a net induced voltage only if the magnetic field varies rapidly with position in space.

Air capacitor microphones require a source of polarization voltage as well as a dc power source for operating the source follower that handles the microphone signal. Electret capacitor microphones are self-polarized but still require power for the source follower signal circuitry. This power is usually supplied by the microphone mixer via the cable connecting the microphone to the mixer. The circuitry employed for accomplishing this must maintain balance of the microphone signal circuitry. Dc circuits that perform this task are called phantom power supplies. One such arrangement is depicted in Fig. 17-27.

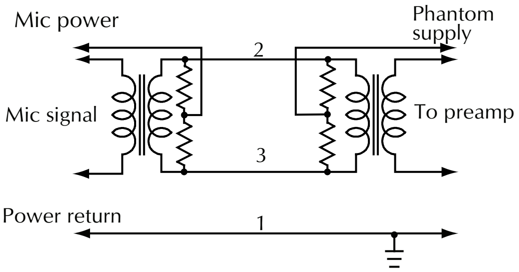


Figure 17-27. Phantom power arrangement for capacitor microphones.

The arrangement of Fig. 17-27 features an output transformer internal to the microphone housing as well as an input transformer internal to the mixer. The dc voltage is applied equally to the microphone signal conductors at pins 2 and 3. The dc return circuit is through the shield on the microphone cable at pin 1. Conductors 2 and 3 have the same dc potential and hence there is no direct current in the transformer windings. In order to accomplish this the resistors denoted as *R* must be carefully matched to be equal to within ±0.1%. This precision is required not only for dc balance but also to maintain a large common mode rejection ratio. Commonly encountered voltage and resistor values are listed below.

Supply Voltage	Resistor Value
12 V	680 Ω ± 0.1%
24 V	1200 Ω ± 0.1%
48 V	6800 Ω ± 0.1%

There is a trend by some designers to employ electronically balanced inputs in the mixer input microphone circuitry. In such instances blocking capacitors must be employed to isolate the differential mixer input from dc while maintaining continuity for the microphone signal. Such an arrangement appears in Fig. 17-28.

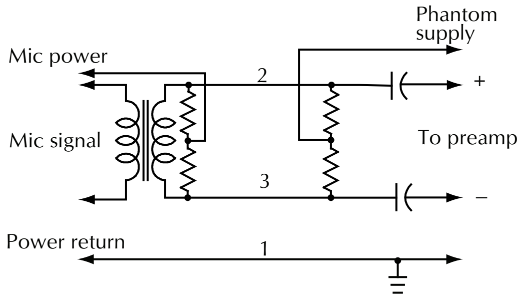


Figure 17-28. Phantom power circuit when electronically balanced inputs are employed.

The phantom power circuits of Figs. 17-27 and 17-28 work well but both have an undesirable feature. The necessity of the employment of matched balancing resistors in both instances limits the current that may be supplied to power the microphone circuitry. This limitation can be removed through the employment of transformers that are center tapped on the appropriate windings. Such transformers would be quite expensive because of the necessity of very accurately having both an equal number of turns on either side of the center tap as well as exact resistance of the turns on either side of the center tap. If this is not accomplished direct current will exist in the transformer winding and the signal circuit will no longer be exactly balanced.

Finally, a word of caution is in order. Sound systems may employ just a few or a very large number of microphone cables not only for microphones but also for link circuits. It is important to maintain correct signal polarity in all microphones, microphone cables, link circuits, all processing electronics, loudspeaker wiring, and loudspeakers.

There are convenient commercial devices called polarity checkers that can be employed to check individual microphones, cables, and overall system polarity. An investment in such devices is modest, time saving, and will earn its keep many times over.

17.9 Measurement Microphones

A collection of measurement microphones whether residing in sound level meters or stand alone devices is an absolute necessity for both sound system installers as well as acoustical consultants. Such a collection must also be supported by an appropriate microphone calibrator system that consists of both the calibrator itself as well as a set of adapters to accommodate the various individual sizes of the microphones in the collection.

For many years there were only two suppliers of quality measurement microphones. Brüel and Kjaer, a Danish firm, and GenRad, a domestic firm. Brüel and Kjaer still exists though not under the original ownership while GenRad no longer exists. Fortunately there are now several new domestic suppliers of quality measurement microphones.

Measurement microphones are dominantly air capacitor or electret capacitor microphones while ceramic piezoelectric units may still be encountered. The standard sizes in terms of capsule diameter are 1 inch, ½ inch, ¼ inch, and ⅛ inch. The larger units have higher sensitivity and lower noise floors. The 1 inch unit is favored for making measurements in quiet environments at frequencies below about 8 kHz. The ½ inch unit is a general purpose one but has high frequency limitations.

Broad frequency band measurements usually require the ¼ inch or ⅛ inch variety particularly if high sound levels are to be encountered. All sizes can have low frequency responses that extend almost to 0 Hz with 3 Hz to 5 Hz being typical with even lower values being possible. A slow leak for allowing the capsule's rear chamber pressure to follow weather induced atmospheric pressure variations determines the low frequency limit.

The geometry of a measurement microphone's physical structure is that of a cylinder with the central axis of the cylinder being perpendicular to the plane that contains the microphone capsule's circular diaphragm. This central axis serves as a reference direction for sound incident on the microphone. Direct sound arrives at 0° relative to this axis while grazing incidence occurs at 90° as illustrated in Fig. 17-29.

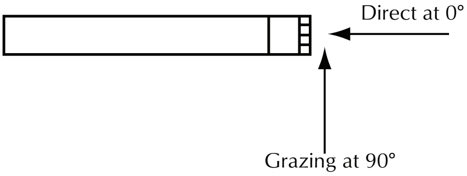


Figure 17-29. Illustration of direct and grazing sound incidence.

Any measurement microphone should be encased in such a fashion that the microphone's physical structure disturbs the sound field in which it is immersed to a minimum degree. When the microphone capsules are smaller than ½ inch in diameter it is impossible to incorporate the necessary circuitry and connector in a uniform cylinder having a diameter equal to that of the capsule. In such instances it is necessary to enclose the circuitry and connector in a larger cylinder that is joined to the capsule by a smoothly tapered section matching the larger diameter to the smaller diameter. A notable example of this is displayed in Fig. 17-30.

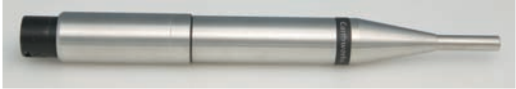


Figure 17-30. An example of a well-engineered tapered microphone structure. (*Photo courtesy of Alex Khenkin of Earthworks, Inc.*)

Measurement Microphone Types

In spite of the smoothness of the microphone enclosure one can not escape the fact that at high frequencies the microphone capsule diameter, d, is comparable to the sound wavelength, λ. When this occurs, the sound field is disturbed by both reflection from the capsule's diaphragm as well as diffraction by the capsule's protective grid and the microphone housing. The degree of this disturbance depends on the angle of incidence of the sound and is greatest for direct incidence.

The acoustic pressure at the diaphragm for directly incident sound at high frequencies can in fact exceed by several decibels that which would have existed in the free field. The free field pressure is that which would have existed if the obstacle presented by the microphone had not been present. It is desirable in such instances to structure the microphone's direct field response such that it is proportional to the free field pressure over as wide a frequency range as possible. It is possible to do so by properly choosing the diaphragm's mechanical resonance frequency and the degree of damping of the mechanical resonance when the microphone capsule is designed.

Measurement microphones designed in this fashion are referred to as being direct or free field microphones. They are valuable for measurements close to sound sources in any environment where direct sound dominates or for measurements at any distance from sound sources in unenclosed spaces.

Measurements made in reverberant environments require a measurement microphone having a different set of characteristics. Here the emphasis is on a microphone response that is flat over a wide frequency range for sound that impinges on the microphone from all directions. Such microphones are referred to as random incidence microphones.

As one would suspect, microphone capsule mechanical structures are different for free field and random incidence microphones. Oftentimes, however, measurements must be made in reverberant environments that also contain directive sound sources. When this is the case, the best results are obtained when the random incidence microphone is oriented such that the direct sound from the source makes an angle of 70° with the microphone axis. Response characteristics of a family of both free field and random incidence measurement microphones are displayed in Fig. 17-31.

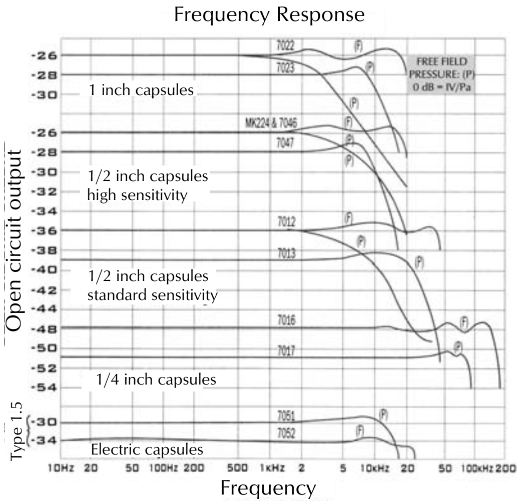


Figure 17-31. A comparison of response characteristics of a family of both free field and random incidence measurement microphones versus capsule sizes. (*Data courtesy of Noland Lewis of ACO Pacific, Inc.*)

In viewing Fig. 17-31 for both the 1 inch and ½ inch capsules, the upper curve in each instance is associated with a capsule optimized for free field measurement. Note also that the upper curve has two branches. One branch is labeled F for free field and corresponds to direct or 0° incidence. The other branch is labeled P and is the pressure response of the same capsule as measured by an electrostatic actuator as described under the next topic. The lower curve in each case is the pressure response of a capsule optimized for random incidence. It should also be noted that the flattest response over the entire audio band is associated with the smallest capsule size.

17.10 Microphone Calibrator

There are several very precise and accurate laboratory techniques for determining a measurement microphone's pressure response as a function of frequency. Manufacturers routinely employ one or more of these techniques to supply a measured calibration curve for each of their measurement microphone capsules. The most often employed technique for laboratory calibration is that of electrostatic actuation. In this technique the microphone capsule is subjected to a combination of static and time varying electric fields. This combination of fields exerts a uniform force over the surface of the diaphragm and in effect allows a measurement of the capsule's pressure response over any frequency range desired.

All laboratory calibration techniques usually require the use of several different instruments. These instruments in turn must periodically be subjected to a calibration process. It is impractical to attempt to employ these laboratory techniques for calibration purposes in the field. Fortunately measurement microphone capsules are quite stable and will maintain their characteristics over long periods of time unless subjected to abuse.

Measurement microphone capsules of different types and sizes can usually be used interchangeably with the same preamplifier provided that differing capsule sensitivities are properly accounted for. This can be done quite readily if one has at hand a device that can subject a given microphone to a known acoustic pressure to within a few percent. Extreme accuracy is not required. For example, a pressure level uncertainty as small as ±0.5 dB corresponds to an acoustic pressure accuracy of ±5.9%.

Acoustic calibrators suitable for field use are based on a relatively simple equation from linear acoustics that relates the acoustic pressure to the change of air density,

$$p = \gamma P_0 \frac{\rho - \rho_0}{\rho_0} \qquad (17\text{-}52)$$

where,

p = acoustic pressure (pascals),

γ = 1.402 (dimensionless),

P_0 = static atmospheric pressure (pascals),

ρ = density of disturbed air (kgm per m^3),

ρ_0 = static air density (kgm/m^3).

The last factor in Eq. 17-52 is the change in air density as a result of compression or expansion divided by the undisturbed air density. This quantity is called the condensation and is positive when the air is compressed and negative when the air is allowed to expand. The compressions and expansions that occur in sound waves in free air are described as being adiabatic. This means that during the compression or expansion heat energy is neither added nor subtracted from the air and the air temperature undergoes changes. The adiabatic process for sound waves results from the fact that air itself is a relatively poor thermal conductor. If air were a good thermal conductor the process would become isothermal. Eq. 17-52 has the same form for an isothermal process except the constant γ would be replaced by 1.

We will now pursue Eq. 17-52 and see what it suggests. The only variable at our disposal in this equation is the disturbed air density ρ. Suppose we have a small container that is filled with air at normal atmospheric pressure. Let the interior volume of this container be V_0. Further suppose that we can introduce some mechanism that periodically can vary the volume of the container above and below V_0 by a small amount ΔV, such that at any instant the volume of the container is $V = V_0 + \Delta V$. If we let the mass of air in the container be represented by m, then the condensation term of Eq. 17-52 can be written as

$$\frac{\rho - \rho_0}{\rho_0} = \frac{\frac{m}{V} - \frac{m}{V_0}}{\frac{m}{V_0}} = \frac{\frac{1}{V} - \frac{1}{V_0}}{\frac{1}{V_0}} = \frac{V_0 - V}{V} \tag{17-53}$$

In Eq. 17-53 the density has been represented by the mass divided by the appropriate volume followed by algebraic simplification. Taking the simplification one step further,

$$\frac{V_0 - V}{V} = \frac{V_0 - (V_0 + \Delta V)}{V_0 + \Delta V} \approx \frac{-\Delta V}{V_0} \tag{17-54}$$

The last step can be justified by simply requiring that the magnitude of ΔV be small as compared with V_0. The conclusion to this point then becomes

$$p \approx -\gamma P_0 \frac{\Delta V}{V_0} \tag{17-55}$$

Eq. 17-55 suggests the following possible structure for a practical microphone calibrator for use in the field.

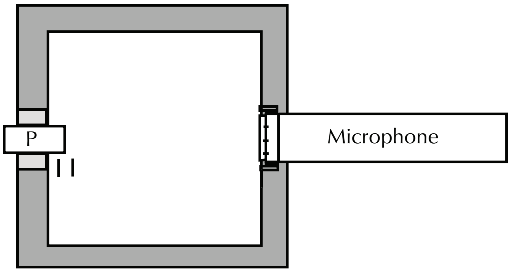


Figure 17-32. Possible microphone calibrator structure.

Fig. 17-32 depicts a rigid enclosure that is fitted with the subject microphone on the right and a bushing mounted movable piston on the left. A motor actuated mechanical linkage, the details of which are not shown in the figure, drives the piston. The nature of the drive is such that the piston is forced to undergo a linear displacement between the indicated limits with the displacement at any instance being a sinusoidal function of time. The piston is depicted at its equilibrium position at which the total enclosed volume of air is V_0.

Let the piston have a cross-sectional area A and let its displacement from equilibrium at any instant be given by $x = x_m \cos(2\pi ft)$. This being the case, the change of volume at any instant will be in the form of $\Delta V = Ax = -Ax_m \cos(2\pi ft)$. The algebraic sign is negative because a positive displacement of the piston decreases the available volume. The acoustic pressure in the enclosure will be everywhere the same as long as the linear dimensions of the enclosure are small as compared with the sound wavelength corresponding to the operating frequency f. This being the case, the acoustic pressure experienced by the microphone's diaphragm will be

$$p = \gamma P_0 \frac{Ax_m \cos(2\pi ft)}{V_0} \tag{17-56}$$

This structure of microphone calibrator is called a piston phone. A variation on this structure replaces the piston and its mechanical drive with a precisely made small loudspeaker powered by an electrical oscillator. For a given piston phone, the piston area, the maximum piston displacement, the equilibrium value of enclosed volume, and the operating frequency are determined at the time of manufacture. The root mean square value of the acoustic pressure produced by such a unit can then be written as

$$\begin{aligned} p_{rms} &= \frac{\gamma P_0}{\sqrt{2}}\left(\frac{Ax_m}{V_0}\right) \\ &= CP_0 \end{aligned} \tag{17-57}$$

where,
C = a constant for the device.

In the foregoing, it has been assumed that the air in the enclosure is operating under true adiabatic conditions. A correction to this assumption must be considered. The walls of the piston phone must be rigid and usually are made of metal. Even though a body of air has poor thermal conductivity, some air is in contact with the walls of the enclosure where heat transfer can occur because of the high thermal conductivity of the metal walls. In accounting for this a small correction factor must be applied to Eq. 17-57 as the true acoustic pressure under this condition is slightly less than predicted. The size of this correction factor hinges on the ratio of the interior volume to the wall surface area. Larger such ratios have smaller correction factors usually only a fraction of 1%. The effect of including this correction factor simply slightly reduces the size of the constant in Eq. 17-57.

The acoustic pressure produced by the device is directly proportional to the ambient atmospheric pressure. For precise work it is necessary to measure the local atmospheric pressure at the time of use. Calibrators are usually supplied with conversion tables for converting barometric atmospheric pressure expressed in mm of Hg to pascals. True piston phones operate at frequencies of a few hundred Hz.

Loudspeaker versions operate at 1000 Hz for convenience in calibrating the A, B, and C scales of sound level meters. All three scales should produce identical readings at 1000 Hz. Typical pressure levels produced by field calibrators are 94 dB, 114 dB, or 124 dB. The loudspeaker versions of calibrators can offer one or more of the above levels as well as operate at frequencies other than 1000 Hz.

Fig. 17-32 was drawn for the largest diameter microphone to be accommodated by the calibrator. Smaller microphone diameters can be accepted when an adapter sized to maintain the interior volume is employed.

Bibliography

Glen M. Ballou. *Handbook for Sound Engineers*, 4th ed. Boston: Focal Press, 2008.

Leo L. Beranek. *Acoustics*, New York: Mc-Graw-Hill, 1954.

Brüel and Kjaer Technical Review. *Measuring Microphones.* Naerum: 1972.

John Eargle. *The Microphone Book*, 2nd ed. Boston: Focal Press, 2004.

P. M. Morse, *Vibration and Sound*, 2nd ed. New York: Mc-Graw-Hill, 1948.

Chapter 18

Loudspeakers and Loudspeaker Arrays

by Eugene Patronis, Jr.

18.1 Loudspeaker Types . . . 347
  Direct Radiators . . . 347
  Direct Radiator Spatial Response . . . 353
  Far Field Directivity of a Direct Radiator . . . 354
  Loudspeaker Cone Behavior . . . 356
  Dome Radiators . . . 358
18.2 Radiated Power . . . 358
18.3 Axial Sound Pressure Level . . . 363
18.4 Efficiency . . . 363
18.5 Loudspeaker Electrical Impedance . . . 364
18.6 Loudspeaker Directivity Factor . . . 365
18.7 Loudspeaker Sensitivity . . . 365
18.8 Direct Radiator Example Calculations . . . 366
18.9 Horns and Compression Drivers . . . 368
  Conical Horn . . . 370
  Cylindrical Horn . . . 371
  Plane Wave Tubes . . . 371
  Catenoidal Horn . . . 372
  Exponential Horn . . . 373
  Constant Directivity Horns . . . 374
18.10 Practical Considerations Involving Horns . . . 374
18.11 Horn Compression Drivers . . . 376
18.12 Crossover Networks . . . 378
  Origin . . . 378
  Electric Filters . . . 379
  Transfer Functions . . . 379
  Higher Order Networks . . . 383
  Three-Way and Higher Crossover Networks . . . 387
  Synthesized Crossover Networks . . . 390
18.13 Loudspeaker Arrays . . . 392
  Acoustic Order of Choices of System Structure . . . 392
  Design Process . . . 393
  Arraying Techniques . . . 395
  Increase Coverage Angle . . . 395
  Increase of Level on Axis . . . 395
18.14 Bessel Array . . . 398
18.15 Line Arrays . . . 400
  Line Array of Discrete Elements . . . 404
  Processed Line Arrays . . . 405
  Distributed Systems . . . 409
  Hybrid Arrays . . . 411
  Split Identical Sources . . . 411
18.16 Vented Enclosure Bass Loudspeakers . . . 412
  Equal Slope Bandpass Subwoofer . . . 416
  Unequal Slope Bandpass Subwoofer . . . 418
18.17 Large Signal Behavior of Loudspeakers . . . 420
18.18 Bibliography . . . 420

A loudspeaker, with the possible exception of the listener's ear, is the final transducer in the sound system chain. The burdens placed on a loudspeaker are both numerous and demanding. The primary burden is to faithfully convert the electrical energy supplied by an amplifier into acoustical energy. Once the conversion is accomplished, the further burden is to direct the acoustical energy through the medium of air to the listening audience while the audience itself may be widely dispersed. Furthermore, the acoustical energy must be directed in such a fashion that the acoustic pressure experienced by all listeners is reasonably uniform. A single device seldom, if ever, can shoulder these burdens. The human audible spectrum encompasses a range of at least ten octaves from 20 Hz to 20 kHz. In the case of sinusoidal excitation, the simple form of the fundamental relationship determining the acoustic pressure in a spherical wave produced by a transducer is written as the proportionality:

$$p \propto \rho_0 \omega S u \tag{18-1}$$

where,
p is the acoustic pressure,
ρ_0 is the undisturbed density of air,
ω is the angular frequency,
S is the effective radiating surface area,
u is the surface velocity.

The product Su is called the volume velocity. The surface velocity in turn is the product of the surface displacement with the angular frequency hence, the relationship might well be written as

$$p \propto \rho_0 \omega^2 S \xi \tag{18-2}$$

where,
ξ is the surface displacement.

At low frequencies where ω is small, a significant acoustic pressure will require a large product of $S\xi$. The displacement itself is limited by mechanical constraints, which requires an even larger surface area and possibly leads to the necessity for multiple units acting in concert. Large surface areas are attended by significant mass.

At high frequencies where ω is large, the acceleration which is the product $\omega^2\xi$ becomes quite large. As a result, the required net driving force, which is the product of the moving mass with the acceleration, becomes excessive. The required force may be made reasonable by a reduction in the moving mass which can be accomplished by a reduction in S. One is thus faced with diametrically opposed constraints in that low frequencies require large S and high frequencies require small S in order to achieve a given appreciable acoustic pressure. In addition there are the questions of uniformity of acoustic pressure with regard to frequency as well as uniformity of radiation pattern with frequency. Taken together, these considerations require that the audible spectrum be divided into two or more bands with appropriately sized transducers operating in the individual bands. This introduces the additional complication of designing the appropriate electrical filters for insertion in the signal paths of the various transducers and arranging the various transducers in such a way that their acoustical outputs combine seamlessly. This is a formidable task indeed and more often than not is only approximately satisfied in practice.

18.1 Loudspeaker Types

Loudspeakers may be broadly classified as being either direct radiators or horns. Further classification can be made with regard to moving surface or diaphragm structure and the physical nature of the driving force. Additionally, a loudspeaker, as distinguished from a loudspeaker system, is often classified as being a low, mid, or high frequency device depending on the portion of the audible spectrum in which its design parameters allow it to excel.

Direct Radiators

The most often employed direct radiators in sound reinforcement systems are based on the pioneering work of Chester W. Rice and Edward W. Kellogg. A tribute to the thoroughness and ingenuity of this work is evidenced by the fact that even though it was first published in 1925, only minor improvements and modifications have been made to the basic design up to the present day. In their initial research, Rice and Kellogg studied a wide range of possible devices for use as a loudspeaker. The devices examined included among others a thermophone, an electrostatic panel, a modulated air stream device, small diaphragm moving coil instruments, and an induction device.

The best results were obtained with an electrostatic panel and a moving coil device having a light paper cone diaphragm mounted on a flat baffle. They found that the electrostatic panel produced quality sound but insulation problems and the necessity for large panels in order to produce significant low frequency output forced their abandonment in favor of the moving coil devices. A diaphragm in the shape of a shallow cone was employed in the preferred moving coil device because this shape was found to offer the most rigidity for the light diaphragm materials employed.

The modern moving coil or electrodynamic loudspeaker is changed little from the original versions of Rice and Kellogg. The modern versions employ permanent magnets rather than the electromagnets in the originals. Modern technology offers a much wider range of cone materials and modern adhesives improve long term longevity even under conditions of hard usage. Additionally, there is now a well-developed theory of enclosure design which was not available to Rice and Kellogg.

Fig. 18-1 is illustrative of a typical modern electrodynamic loudspeaker.

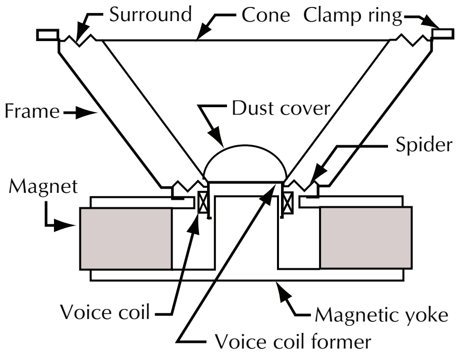


Figure 18-1. Loudspeaker structure.

The parameters necessary to characterize the small signal behavior of this form of direct radiator are partly electrical, partly mechanical, and partly acoustical. The mechanical parameters are denoted as M, K, S, and R_m. M is the effective moving mass of the cone or diaphragm. It includes all of the cone mass as well as that of the voice coil and the voice coil former with some additional fractional contribution from the centering spider and the outer surround. K is the stiffness constant of the suspension and is constituted of contributions from both the centering spider and the surround. S is the effective radiating area. This is not the actual surface area of the cone, which is a larger quantity, but rather, is the area of a circle whose diameter is slightly larger than the cone diameter. It is slightly larger than the actual cone diameter because the inner portion of the surround moves with the cone and contributes to the radiating surface. R_m is the mechanical resistance associated with the suspension elements including both the surround and the spider.

The electrical parameters are denoted as being R_e, L_e, B, and l. R_e is the direct current resistance of the voice coil, L_e is the self inductance of the voice coil, B is the radially directed magnetic induction in the magnetic air gap in which the voice coil resides, and l is the total length of voice coil conductor residing in the magnetic air gap. The acoustical parameters are denoted as R_r and X_r. R_r is a frequency dependent function called the radiation resistance. In addition to its frequency dependence, the radiation resistance depends also on the radiating surface area S and on boundaries in the vicinity of the radiator. The radiation resistance is quite small at low frequencies and as a consequence, has little influence on the motion of the radiating surface in the low frequency limit.

X_r is also a frequency dependent function. Its value at any frequency depends on the radiating surface area and the radiator boundaries. Its principal physical effect at low frequencies is to act as if a small additional mass is attached to the cone or diaphragm. In other words, at low frequencies it increases the cone's inertia. Values of the electrical and mechanical parameters that are typical for a mid range direct radiator are listed in Table 18-1.

Table 18-1. Electrical and Mechanical Parameters of a Typical Mid Range Direct Radiator.

Parameter Description	Symbol and Dimensions	Typical Values
Radiator radius	a (meter)	0.10
Radiator area	$S = \pi a^2$ (meter^2)	0.0314
Moving mass	M (kilogram)	0.03
Suspension stiffness	K (Newton-meter^{-1})	3000.0
Suspension resistance	R_m (kilogram-s^{-1})	2.5
Voice coil resistance	R_e (Ohm)	6.0
Voice coil inductance	L_e (Henry)	0.001
Magnetic induction in gap	B (Tesla)	1.0
Length of conductor in gap	l (meter)	10.0

It should be noted that the values listed in Table 18-1 all employ SI units. These are the units currently employed in all scientific work. In the United States, the past common practice was to express loudspeaker sizes in terms of a diameter expressed in inches with the common sizes being 2, 4, 6, 8, 10, 12, 15, and 18 inches. In many other countries the practice was to express the diameter in millimeters. In any event, the stated sizes are misleading in that hardly any feature of the device corresponds to the stated size. Fortunately there is a rule of thumb which is accurate to within about 5%. This rule is that the stated diameter in inches is approximately equal to the effective radius of the radiator expressed in centimeters. For example, the parameters of Table 18-1 are those of a 10 in diameter driver. The rule would say that the effective radius of this device is then 10 cm. Ten cm correspond to 0.1 m and that is indeed the effective radiator radius.

Regardless of whether one is dealing with a low, mid, or high frequency direct radiator, at the low end of the respective operating range, the device is considered to act as a flat piston for the purpose of analyzing the dynamics and acoustics of the radiator. This is justified because at the lower frequencies the cone undergoes axial motion as a unit and the wavelengths are larger than the dimensions of the radiator. If such a radiator is suspended in free air devoid of any baffle or back enclosure it would attempt to radiate from its rear side as well as its front side. In this circumstance when the piston moves so as to compress the air in front of it, the air to the rear will be rarefied, meaning that the front side and the rear side are out of polarity. Such an arrangement constitutes what is called an acoustic dipole and is a poor radiator particularly at low frequencies.

It is necessary to separate acoustically the front side from the back side either by means of a large flat baffle or by enclosing the rear side. The latter option is more convenient in practice even though the back enclosure itself introduces further complications. If one can manage to keep the back enclosure dimensions small compared with any wavelength in the pass band of the transducer, the air trapped in an unvented back enclosure acts simply as a spring effectively increasing the overall stiffness of the loudspeaker. This additional stiffness associated with the enclosure is given by

$$K_b = \frac{\gamma P_0 S^2}{V_0} \tag{18-3}$$

where,

γ is dimensionless with a value of 1.4 and is the adiabatic constant for air,

P_0 is the static air pressure,

S is the effective radiator surface area,

V_0 is the actual air volume in the interior of the back enclosure.

Clearly, the back enclosure does influence the dynamics of the radiator, as the overall stiffness is now larger than that of the radiator alone. The total stiffness now becomes

$$K' = K + K_b \tag{18-4}$$

When a direct radiator having a rear enclosure of small dimensions is suspended freely in air far from an external bounding surface, it radiates into all of space, into a total solid angle of 4π steradians. On the other hand, if it is positioned on a large plane surface it radiates into a solid angle only half as large, namely 2π steradians. In both instances the radiation load acts only on the exposed face of the radiating piston. This radiation load is characterized by what is termed the radiation impedance. This radiation impedance in the instance of half space radiation (2π solid angle) is given by

$$\begin{aligned} Z_r &= R_r + jX_r \\ &= \rho_0 c S[R(2ka) + jX(2ka)] \end{aligned} \tag{18-5}$$

where,

ρ_0 is the undisturbed density of air,

c is the speed of sound,

S is the effective piston area,

The expression in the brackets is called the piston function.

The piston function is obviously complex. If one lets x represent $2ka$ where k in turn is $2\pi/\lambda$ then the real and imaginary parts of the piston function may be written as

$$\begin{aligned} R(x) &= 1 - \frac{2J_1(x)}{x} \\ &= \left(\frac{x^2}{2 \cdot 4} - \frac{x^4}{2 \cdot 4^2 \cdot 6} + \frac{x^6}{2 \cdot 4^2 \cdot 6^2 \cdot 8} - \ldots\right) \end{aligned} \tag{18-6}$$

$$\begin{aligned} X(x) &= \frac{2H_1(x)}{x} \\ &= \frac{4}{\pi}\left[\frac{x}{3} - \frac{x^3}{3^2 \cdot 5} + \frac{x^5}{3^2 \cdot 5^2 \cdot 7} - \ldots\right] \end{aligned} \tag{18-7}$$

In Eq. 18-6, $J_1(x)$ is the first order Bessel function of the first kind and in Eq. 18-7, $H_1(x)$ is the first order Struve function. The real part of the piston function is displayed graphically in Fig. 18-2 while the behavior of the imaginary part of the piston function is exhibited in Fig. 18-3.

It should be noted that at high frequencies where $2ka$ is large, the piston resistance function approaches unity whereas the piston reactance function approaches zero. On the other hand when $2ka$ is small, one needs only to retain the first term in the infinite series description of either $R(x)$ or $X(x)$. Now if one does this and replaces x by $2ka$ the result becomes

$$R(2ka) \approx \frac{1}{2}(ka)^2 \tag{18-8}$$

and

$$X(2ka) \approx \frac{8ka}{3\pi} \tag{18-9}$$

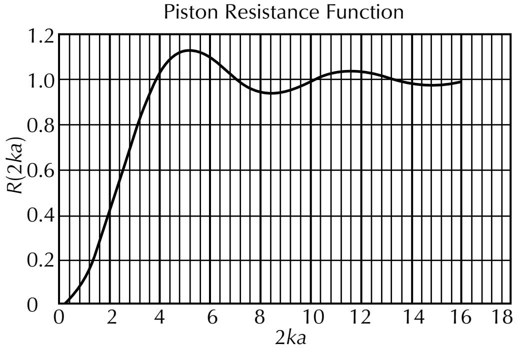


Figure 18-2. Real part of piston function.

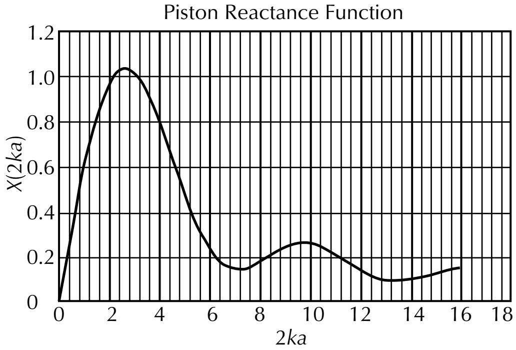


Figure 18-3. Imaginary part of piston function.

It is now possible to obtain a simple expression for the frequency dependence of the radiation impedance in the range of frequencies where the radiated wavelengths are large compared with the dimensions of the piston radiator. In doing so, it is necessary to recognize that

$$\begin{aligned} k &= \frac{2\pi}{\lambda} \\ &= \frac{\omega}{c} \end{aligned} \tag{18-10}$$

where,

λ is the wavelength,

ω is the angular frequency,

c is the speed of sound.

One needs to simply substitute the results of Eqs. 18-8, 18-9, and 18-10 into Eq. 18-5 to obtain

$$Z_r \approx \frac{\rho_0 S^2 \omega^2}{2\pi c} + j\frac{\rho_0 8Sa\omega}{3\pi} \tag{18-11}$$

In the high frequency limit where the wavelength is small compared with the dimensions of the radiator, the corresponding result is

$$Z_r \approx \rho_0 cS \tag{18-12}$$

For intermediate frequencies between the limiting extremes it is necessary to employ the complete expression of Eq. 18-5 where numerical evaluation can be facilitated by extracting values from the graphs of Figs. 18-2 and 18-3. In all cases, if the radiator works into a solid angle of 4π meaning a full space rather than a half space, the radiation impedance values will be halved because the acoustic pressure will be only half as large in this instance.

In order to determine the acoustic pressure produced by a direct radiator, it is first necessary to obtain an expression for the piston velocity. This can be accomplished by solving what is termed the equation of motion of the direct radiator. The equation of motion is obtained by writing an expression for the net force acting on the radiating piston and equating this to the product of the piston mass with the piston acceleration as expressed in Newton's second law of motion. In the general case this leads to a second order differential equation. The labor is greatly reduced by seeking only the steady state solution while employing only a sinusoidal excitation expressed as a complex exponential function of time as employed in Chapter 6 *Mathematics for Audio*. This procedure will allow the conversion of the differential equation into an algebraic equation with the piston velocity being the dependent variable. Under these conditions, the general equation for the piston velocity becomes

$$u = \frac{\dfrac{vBl}{R_e + j\omega L_e}}{Z_m + \dfrac{B^2 l^2}{R_e + j\omega L_e}} \tag{18-13}$$

where,

u is the piston velocity expressed as a complex exponential function of time, i.e., as a phasor,

v is a phasor representing the voltage applied across the voice coil terminals of the radiator,

Z_m is the mechanical impedance of the radiator which is given by

$$Z_m = R_m + R_r(2ka) + j\left[\omega M + X_r(2ka) - \frac{K'}{\omega}\right] \tag{18-14}$$

Practitioners are often faced with the problem of determining the optimum volume of a simple back enclosure for a direct radiator. The pursuit of a solution to this problem is facilitated by examining the simplifications that occur in Eqs. 18-13 and 18-14 when the operating angular frequency is small. For this circumstance, ωL_e is considerably smaller than R_e and Eq. 18-13 simplifies to become

$$u = \frac{\dfrac{vBl}{R_e}}{Z_m + \dfrac{B^2 l^2}{R_e}} \tag{18-15}$$

Additionally, when ω is small, the radiation impedance takes on the form of Eq. 18-11 and the mechanical impedance can now be written as

$$Z_m = R_m + \frac{\rho_0 S^2 \omega^2}{2\pi c} + j\left[\omega\left(M + \frac{\rho_0 8Sa}{3\pi}\right) - \frac{K'}{\omega}\right] \tag{18-16}$$

In examining Eq. 18-16 it should be noted that at low frequencies the effect of the radiation reactance is to increase the overall mass to a new value M' with

$$M' = M + \frac{\rho_0 8SA}{3\pi} \tag{18-17}$$

It is also true that for small ω the radiation resistance is small when compared with the mechanical resistance so Eq. 18-15 now becomes

$$u = \frac{\dfrac{vBl}{R_e}}{R_m + \dfrac{B^2 l^2}{R_e} + j\left[\omega M' - \dfrac{K'}{\omega}\right]} \tag{18-18}$$

Eq. 18-18 exhibits a resonance behavior in that there exists a particular value of ω called ω_0 such that the denominator of Eq. 18-18 takes on a minimum value and the amplitude of u takes on a maximum value. Additionally, when $\omega = \omega_0$, the denominator is purely real and the piston velocity is in phase with the driving voltage. This will occur when

$$\omega_0 M' - \frac{K'}{\omega_0} = 0 \tag{18-19}$$

Upon solving Eq. 18-19 we obtain

$$\omega_0 = \sqrt{\frac{K'}{M'}} \tag{18-20}$$

As is usual the resonance frequency as distinguished from the resonance angular frequency is given by

$$\begin{aligned} f_0 &= \frac{\omega_0}{2\pi} \\ &= \frac{1}{2\pi}\sqrt{\frac{K'}{M'}} \end{aligned} \tag{18-21}$$

In describing a resonance, in addition to a specification of the resonance frequency or angular frequency, one needs to give an indication of whether the resonance is well defined and sharp or whether it is poorly defined and broad. This is done by a specification of the quality factor or Q of the resonance. The quality factor in this instance is referred to as being the total quality factor and is defined to be

$$Q_t \equiv \frac{\omega_0 M'}{R_m + \dfrac{B^2 l^2}{R_e}} \tag{18-22}$$

Now if one were to multiply both the numerator and denominator of Eq. 18-18 by the quantity $j\omega/\omega_0{}^2 M'$, rearrange terms and identify Q_t where it appears, the result will appear as

$$u \approx \frac{\dfrac{j\omega vBl}{\omega_0{}^2 M' R_e}}{1 - \dfrac{\omega^2}{\omega_0{}^2} + \dfrac{j\omega}{Q_t \omega_0}} \tag{18-23}$$

Upon having the equation for the piston velocity in hand, it is now possible to write the expression for the acoustic pressure produced on the axis of the direct radiator for distances much greater than the radius a of the radiator. In doing so, it also must be remembered that this result is valid only at low frequencies where the wavelength is very much larger than the radius of the piston.

$$p \approx \left[\frac{V_m Bl \rho_0 S}{2\pi r M' R_e}\right]\left[\frac{\dfrac{-\omega^2}{\omega_0{}^2}}{1 + \dfrac{j\omega}{Q_t \omega_0} - \dfrac{\omega^2}{\omega_0{}^2}}\right]\left[e^{j(\omega t - kr)}\right] \tag{18-24}$$

Eq. 18-24 contains the product of three bracketed terms. The bracketed term on the left contains the essential amplitude factors among which is the amplitude of the applied driving voltage denoted by

V_m and the axial radial distance between the observer and the source which is denoted by r. The bracketed term on the right describes the wave propagation. The middle bracketed complex term embodies the frequency response of the radiator including both amplitude and phase information. The amplitude response of the radiator is given by the magnitude of the complex quantity in the bracketed term in the middle. This magnitude is given by

$$Mag = \frac{\dfrac{\omega^2}{\omega_0^2}}{\sqrt{\left(1 - \dfrac{\omega^2}{\omega_0^2}\right)^2 + \dfrac{\omega^2}{Q_t^2 \omega_0^2}}} \tag{18-25}$$

The smoothest and most flat response is termed the maximally flat response and occurs when Q_t is assigned the value $1/(\sqrt{2})$. The behavior of Eq. 18-25 is exhibited in Fig. 18-4 for three choices of Q_t, including the optimum choice that yields the maximally flat behavior. The plots are presented in the form 20 dBlog(Mag) versus log(x) where $x = \omega/\omega_0$.

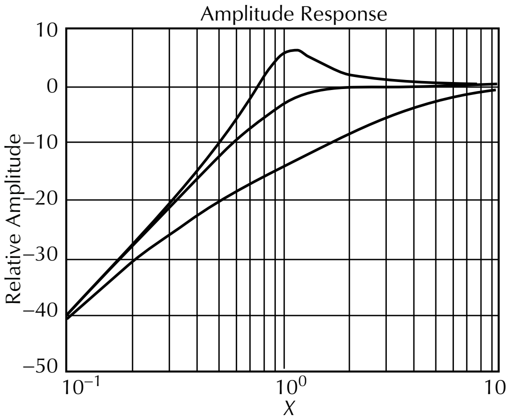


Figure 18-4. Relative amplitude response for $Q_t = 2.0$ (upper curve), $Q_t = 0.707$ (middle curve), and $Q_t = 0.2$ (lower curve).

Upon imposing the condition of optimum response, one learns that

$$\begin{aligned} Q_t &= \frac{1}{\sqrt{2}} \\ &= 0.7071 \\ &= \frac{\omega_0 M'}{R_m + \dfrac{B^2 l^2}{R_e}} \\ &= \frac{\sqrt{K'M'}}{R_m + \dfrac{B^2 l^2}{R_e}} \end{aligned} \tag{18-26}$$

Turn now to the situation that exists when the direct radiator operates without a back enclosure and is far from any boundary. Both faces of the radiator are now operating in free air and each radiates into a 4π solid angle. In this instance, there is a radiation impedance acting on each face the value of which is one-half of the former value. There are now two such impedances so that the total radiation impedance is the same as that for the back enclosed device positioned on a large plane boundary. The overall stiffness, however, is just that of the suspension. The resonance angular frequency is now denoted as the free air value and is given by

$$\omega_{fa} = \sqrt{\frac{K}{M'}} \tag{18-27}$$

Similarly, the quality factor for the free air resonance appears as

$$Q_{fa} = \frac{\sqrt{KM'}}{R_m + \dfrac{B^2 l^2}{R_e}} \tag{18-28}$$

It must be noted that

$$\omega_0 = \sqrt{\frac{K'}{K}}\,\omega_{fa} \tag{18-29}$$

and

$$Q_t = \sqrt{\frac{K'}{K}}\,Q_{fa} \tag{18-30}$$

Two conclusions can be drawn immediately from Eqs. 18-29 and 18-30. The resonance with a back enclosure is higher than that of free air and in order for a direct radiator to be a candidate for application with an enclosed back, its quality factor when operating in free air must be less than 0.7071. These consequences result because in both instances the multiplying factor is greater than one. Upon setting $Q_t = 0.7071$ and substituting for K', Eq. 18-30 becomes

$$0.7071 = \sqrt{\frac{K + \dfrac{\gamma P_0 S^2}{V_0}}{K}}\,Q_{fa} \tag{18-31}$$

Eq. 18-31 can now be solved for the required back enclosure volume to obtain

$$V_0 = \frac{\gamma P_0 S^2}{K\left[\left(\frac{0.7071}{Q_{fa}}\right)^2 - 1\right]} \tag{18-32}$$

There are some techniques other than a simple back enclosure that can be employed to enhance the low frequency performance of a direct radiator. These techniques involve the employment of a vented enclosure or passive radiators. A discussion of these techniques is given in Section 18-16.

In arriving at the expressions for the acoustic pressure in the foregoing analyses it must be remembered that the calculations are based on certain assumptions. The first assumption is that the point of observation of the pressure is on the axis of the radiator at a distance that is large in comparison with the radius of the radiator, actually in what is called the far field. The second assumption is that the frequency range is that for which $2ka$ is less than one. In examining more general cases, one must first explore the general expression for the acoustic pressure on the axis of the direct radiator for half space radiation.

Direct Radiator Spatial Response

Fig. 18-5 represents a piston radiating into a half space. The radius of the piston is a and the overall frontal surface area is S. The piston undergoes sinusoidal motion along the z-axis with a surface velocity described by $u_m e^{j\omega t}$. If the piston were very small such that its radius was small compared to the wavelength for all frequencies for which it is employed, the acoustic pressure at the observation point O could be stated simply as

$$p(r, t) = \frac{j\rho_0 \omega S u_m}{2\pi r} e^{j(\omega t - kr)} \tag{18-33}$$

where,

p is the acoustic pressure,

k is 2π divided by the wavelength,

r is the radial distance measured from the center of the piston,

t is the independent variable time,

ρ_0 is the undisturbed density of air,

ω is the angular frequency,

S is the piston area,

u_m is the piston velocity amplitude.

For the more general case where the piston radius is of the order of or exceeds the wavelength, it is necessary to consider the piston's surface to be made up of an infinitely large number of infinitesimally small elements each of which has an infinitesimal area dS and contributes an infinitesimal pressure dp. Such an element is depicted in the figure by a black dot. This element is a distance ρ from the center of the piston where ρ makes an angle φ with the x-axis. The radial distance from the position of the element to the observation point O is denoted as r'. The radial distance from the center of the piston to the observation point is r and this radial line makes an angle θ with the z-axis. The total acoustic pressure at the observation point is obtained by summing the contributions from all elements taking due account of the fact that the phase of the individual contributions varies because r' varies as one travels over the piston's surface. The general mathematical statement in its simplest form appears as

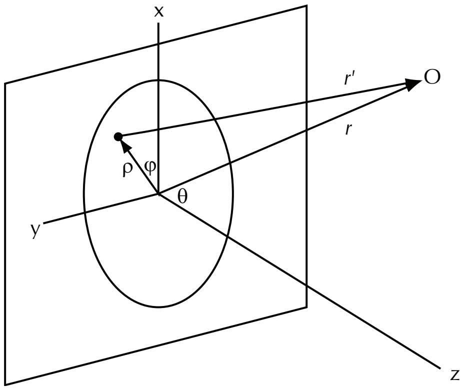


Figure 18-5. Piston radiating into a half space.

$$p(r, \theta, t) = \frac{j\rho_0 \omega u_m}{2\pi} \int_S \frac{1}{r'} e^{j(\omega t - kr')} dS \tag{18-34}$$

The calculation indicated in Eq. 18-34 is extremely difficult to carry out for the general case. Fortunately there are two very useful special cases for which the calculation may be made in closed form. The first of these yields the acoustic pressure at all axial points. This calculation yields the acoustic pressure on the axis that is valid for all axial distances, all piston sizes, and all frequencies. In this calculation the angle θ is always zero. The second special case calculation yields the acoustic pressure for points on or off the axis that is valid for all piston sizes and frequencies but is restricted to observation points which are in the far field. The conditions that constitute the far field will emerge in the course of the examination of the first special case.

When one sets $\theta = 0$, Eq. 18-34 takes on the form

$$p(r,t) = \frac{j\rho_0\omega u_m}{2\pi}\int_0^a\int_0^{2\pi}\frac{e^{\left(-jk\sqrt{r^2+\rho^2}\right)}}{\sqrt{r^2+\rho^2}}d\varphi\rho d\rho \qquad (18\text{-}35)$$

where,

φ is the azimuthal angle in Fig. 18-5.

Eq. 18-35 can be integrated to obtain

$$p(r,t) = \frac{\rho_0\omega u_m}{k}\left\{1 - e^{\left\{-jkr\left(\sqrt{1+\frac{a^2}{r^2}}-1\right)\right\}}\right\}e^{j(\omega t - kr)} \qquad (18\text{-}36)$$

The information with regards to the spatial variation of the acoustic pressure given by Eq. 18-36 is contained in the term contained within the braces. It is instructive to plot the magnitude of this term to obtain the result that is displayed in the following figure.

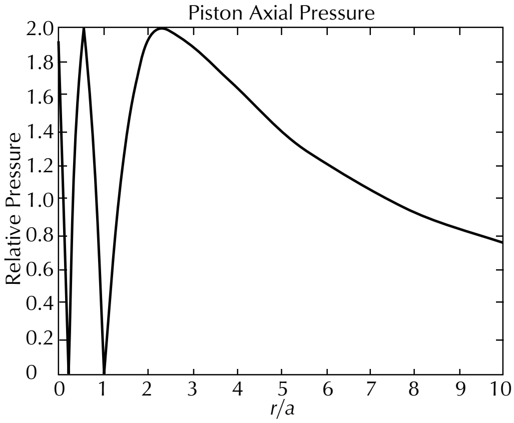


Figure 18-6. Axial distance pressure variation for a piston radiator.

Fig. 18-6 is plotted by assigning *ka* a value of 5π, where *ka* is the circumference of the piston divided by the wavelength. This choice was made in order to clearly delineate three regions of interest. Note that when *r/a* is ten or above the axial pressure variation falls inversely with distance. In this region *r* is much greater than *a* and if in addition *r* is also greater than *S/*λ, then one is in a region that is called the far field. In this region Eq. 18-36 becomes

$$p(r,t) = \frac{j\rho_0\omega S u_m}{2\pi r}e^{j(\omega t - kr)} \qquad (18\text{-}37)$$

On the other hand, if *r* is much smaller than a and if *ka* is considerably smaller than one, then Eq. 18-36 becomes

$$p(r,t) = \frac{j\rho_0\omega S u_m}{\pi a}e^{j\omega t} \qquad (18\text{-}38)$$

where,

a is the piston radius.

Equation 18-38 describes the axial acoustic pressure in what is called the extreme near field. If one divides Eq. 18-37 by Eq. 18-38, one obtains

$$\frac{p_f}{p_n} = \frac{a}{2r}e^{-jkr} \qquad (18\text{-}39)$$

where,

p_f is the far field acoustic pressure calculated by Eq. 18-37,

p_n is the near field acoustic pressure calculated by Eq. 18-38.

This is a very important result that was first pointed out by Keele. The quantity e^{-jkr} is a phase factor whose magnitude is always one, hence Eq. 18-39 implies that the pressure amplitude in the far field can be accurately predicted by a near field pressure amplitude measurement. Furthermore, this prediction is independent of frequency provided that true near field measurement conditions are satisfied. Near field conditions require that *ka* be small compared with one or *ka* should not exceed about 0.1. Eq. 18-39 is the justification for measuring the frequency response of low frequency loudspeakers in the near field. Finally, the region between *r* greater than 0.1*a* and *r* less than 10*a* is called the intermediate field. The pressure maxima and minima in this region result from interference effects of radiations emanating from different spatial zones on the piston surface. This extended region is thus a poor choice of one in which to perform frequency response measurements.

Far Field Directivity of a Direct Radiator

When far field conditions are satisfied, the acoustic pressure on or off of the piston axis is given by the following integral that sums radiation from all portions of the piston face while taking proper account of phase differences.

$$p(r, \theta, t) = \frac{j\rho_0 \omega S u_m}{2\pi r} e^{j(\omega t - kr)} \int_0^{\pi} e^{jka\sin\theta\cos\varphi} \sin^2\varphi d\varphi \qquad (18\text{-}40)$$

This integration can be carried out in closed form resulting in

$$p(r, \theta, t) = \frac{j\rho_0 \omega S u_m}{2\pi r}\left[\frac{2J_1(ka\sin\theta)}{ka\sin\theta}\right] e^{j(\omega t - kr)} \qquad (18\text{-}41)$$

where,
J_1 is the first order Bessel function of the first kind.

When one compares Eq. 18-41 which describes the far field acoustic pressure for arbitrary observation points with Eq. 18-37 which describes the far field acoustic pressure for points on the axis of the radiator, it is found that the two equations differ by the factor

$$\left[\frac{2J_1(ka\sin\theta)}{ka\sin\theta}\right] \qquad (18\text{-}42)$$

This factor is called the directivity function. This function is of the form $2J_1(x)/x$ where $J_1(x)$ is the Bessel function of the first kind and first order. The behavior of Eq. 18-42 is readily explored in two ways. Initially, the value of the function is plotted versus a range of values for the product $ka\sin\theta$. This plot is depicted in Fig. 18-7.

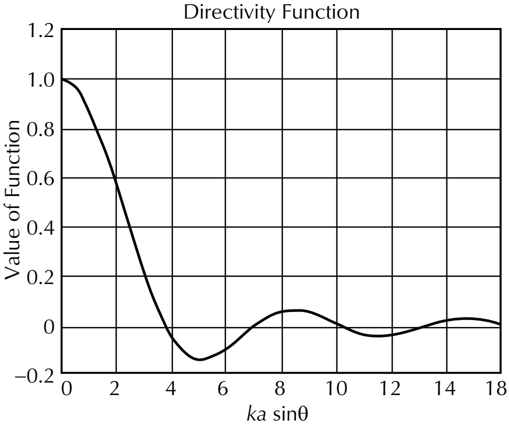


Figure 18-7. Piston far field directivity function.

Fig. 18-7 indicates several points of interest. It should be noted that when the angle θ is zero, which is true for all axial points, the directivity function has the value of unity regardless of the value of *ka*. Additionally, when *ka* is greater than zero, an increase in the value of θ brings about a decrease in the value of the function meaning that the pressure off axis is less than the pressure on axis. This means that the radiator has a directional behavior. Finally, there are regions where the directivity function is negative. A negative value of the directivity function is an indication of a polarity change or a phase shift of π. The directional character of the piston behavior will be explored through a series of polar plots.

Fig. 18-8A displays the polar plot of half space radiation as calculated from Eq. 18-42 with a direct radiator piston radius of 0.1 m (a 10 inch diameter loudspeaker) at a frequency of 125 Hz. The radial value in the figure represents the absolute magnitude of the directivity function on a linear scale. At this low frequency the directivity is essentially independent of the polar angle.

In Fig. 18-8B the operating frequency is now 2 kHz. At this frequency the directivity function is nearly zero when the angle takes on the value of ±π/2 or ±90°. At this frequency, the acoustic pressure for points perpendicular to the piston axis is essentially zero.

As the frequency increases above 2 kHz for this piston, side lobes begin to form. Fig. 18-8C illustrates this behavior where the operating frequency is now 4 kHz. At a frequency of 4 kHz for this piston, there is very nearly closure of the first pair of side lobes.

The operating frequency beyond which side lobes will be formed for a given piston size can be determined by the following considerations. The directivity function first becomes zero when $ka\sin\theta$ takes on the value of 3.83. In order to have at most a single lobe in the radiation pattern, this value must occur when the absolute value of $\sin\theta$ is one, i.e., when θ is ±π/2. Therefore,

$$ka = 3.83$$

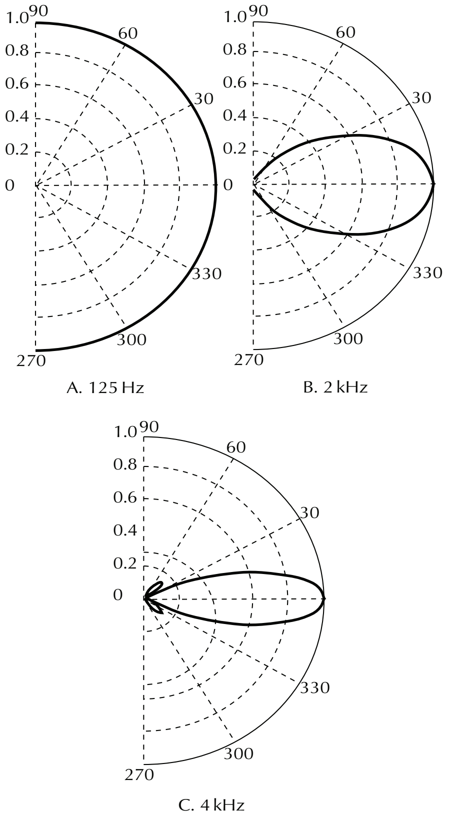
or

$$f = \frac{3.83c}{2\pi a} \qquad (18\text{-}43)$$

As an example, if a piston of 0.1 m radius is to have at most a single lobe when c is 344 m/s, then the maximum operating frequency will be given by

$$f = \frac{3.83 \times 344}{2\pi \times 0.1}$$
$$= 2096 \text{ Hz}$$

This frequency is approximately that for which Fig. 18-8B was drawn.



A. 125 Hz

B. 2 kHz

C. 4 kHz

Figure 18-8. Piston directivity at 125 Hz, 2 kHz, and 4 kHz.

A useful datum for all radiators is the coverage angle. In loudspeaker coverage calculations, the coverage angle is usually defined as the angular interval between the half pressure points. These are the points where the pressure response is –6 dB relative to the pressure on axis.

All piston radiators are non-directional at sufficiently low frequencies and become increasingly directional as the operating frequency increases and eventually arrive at the closure of a single lobe at the maximum frequency calculated according to Eq. 18-43. It is reasonable to inquire as to the coverage angle at this frequency. In order to have just a closed single lobe *ka* must have the value 3.83. The question to be answered becomes what value assigned to $ka\sin\theta$ will cause the directivity function to become 0.5. This value may be extracted by a close examination of Fig. 18-7 or by consulting a tabulation of the directivity function. This value is found to be 2.2. Therefore,

$$\begin{aligned}\theta &= \arcsin\left(\frac{2.2}{3.83}\right)\\ &= 35^\circ\end{aligned}$$

This angle is measured relative to the axis hence the total coverage angle is twice this value or 70º. Table 18-2 relates piston size to the maximum operating frequency that results in closure of a single lobe yielding a minimum coverage angle of 70°.

Table 18-2. Piston Size as Related to Maximum Frequency for Single Lobe Pattern

Approximate Diameter (in)	Piston Radius (m)	Maximum Frequency (kHz)
1.0	0.01	21.0
3.0	0.03	7.0
5.0	0.05	4.2
8.0	0.08	2.6
10.0	0.1	2.1
12.0	0.12	1.75
15.0	0.15	1.4

Although the foregoing is correct for true pistonic motion in which all elements of the piston surface execute axial motion in phase with a common surface velocity, one must question whether actual loudspeaker cones mimic this behavior. This is explored in the next section.

Loudspeaker Cone Behavior

There are several reasons why actual loudspeaker cones have broader radiation patterns than those possessed by true pistons. In some instances, special cone construction techniques are employed which progressively allow the outer portions of the cone to decouple from the inner portions as the cone operating frequency increases.

As an example, a cone of 0.1 m actual radius may have all portions moving as a unit at 200 Hz. As a result of decoupling of an outer portion its effective radius at 2 kHz may be only 0.05 m. By the time 20 kHz is reached, further decoupling may have reduced the effective radius to 0.015 m. This behavior is accomplished by molding the cone with several concentric corrugations or ribs that act as springs or compliances connecting the various intervening cone sections. Furthermore, the nature of wave propagation in the cone material itself can influence the effective radius of the cone. The internal wave in the cone material begins at the voice coil-cone juncture and moves through the material of the cone towards the surround with a speed governed by the nature of the wave and the material properties of the cone. There is thus a phase lag between the radiations emitted by the inner and outer portions of the cone. This phase lag is insignificant at low frequencies but can produce dips in the high frequency response of the cone. This internal wave propagation can also lead to the existence of complicated resonances in the cone. Resonance behaviors in cones are generally referred to as "cone breakup." In order to obtain some understanding of these phenomena one must explore the character of the driving force acting on a typical cone. This situation is depicted in Fig. 18-9.

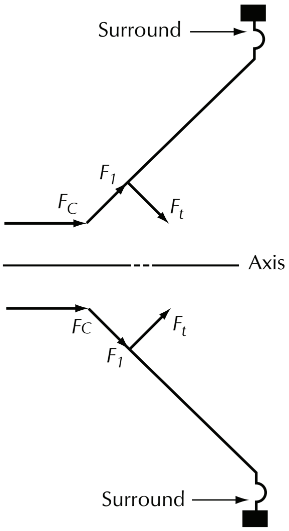


Figure 18-9. Driving force acting on a loudspeaker cone.

F_c in the figure represents the driving force exerted at the truncated apex of the cone by the voice coil. This driving force is resolved into the two components F_l and F_t. F_l is the longitudinal component of the driving force acting parallel to the cone's surface while F_t is the transverse component of the driving force acting perpendicular to the cone's surface. This is better understood by an examination of Fig. 18-10 in which is displayed an enlarged drawing of a small element of the cone attached to the voice coil. This drawing displays not only the driving force acting on the element but also the forces acting on the element produced by the rest of the cone which is beyond the element under consideration.

In Fig. 18-10A, F_l is the longitudinal component of the driving force acting on the element while F_{lc} is the longitudinal component of the force on the element exerted by the rest of the cone beyond the element. The result of these forces is to distort the normal shape of the element, which is shown as being shaded, into a deformed state displayed as white. Along the longitudinal direction this deformation is that of compression and produces a wave motion of the type in which the disturbance is in the direction of propagation just as is the case for a sound wave in air. In this instance, however, the longitudinal wave is in the cone material itself and has a sound speed determined by the properties of the cone material. Additionally, the deformation of the element exerts a stress indicated by the black dot acting normal to the plane of the drawing. This will be considered later.

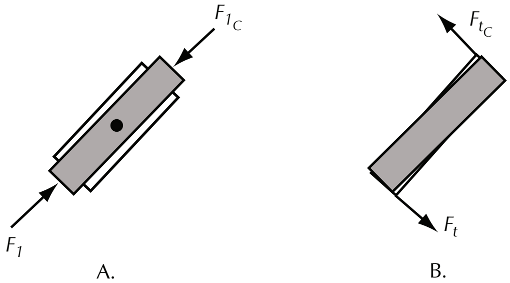


Figure 18-10. Detail of forces acting on a cone element.

Similarly in Fig. 18-10B, F_t is the transverse component of the driving force acting on the element and F_{tc} is the transverse force acting on the element exerted by the rest of the cone beyond the position of the element. These shearing forces deform the shape of the element transverse to the direction of propagation of the disturbance. These composite actions, then, induce both longitudinal and transverse waves traveling in the cone material from the voice coil-cone juncture to the outer surround.

Even though the differential equation describing the propagation of the composite disturbance in the cone is well known, the solution to the equations is not. The mathematical complexity presented by this relatively simple appearing structure is immense. The solutions are strongly influenced by the cone shape, the homogeneity or lack thereof of the cone material, and the boundary conditions that exist at the voice coil and at the surround. Furthermore, the wave which travels in the cone is a damped wave, i.e., its amplitude is attenuated as it travels through the material of the cone and this disturbance, when it arrives at the surround, is subject to reflection at this boundary thus giving rise to standing waves. At certain resonant frequencies the amplitude of these standing waves is quite large and becomes one contributing factor to what is termed cone breakup. These resonant frequencies or normal modes are determined by the cone size as well as shape and are influenced by the properties of the cone material as well as the surround. These resonant standing wave modes are called the concentric modes. This type of undesirable resonant effect is best treated by careful treatment of the terminating impedance that exists at the surround so as to absorb the wave energy of the cone rather than reflect it.

Attention is now returned to the stress indicated by the black dot at A in Fig. 18-10. This stress is called a hoop stress because of the geometrical shape of a cone element that is subjected to this stress. Such an element is formed by slicing the cone by two parallel planes separated by a small distance with both planes being perpendicular to the loudspeaker axis. Hoop stress is the origin of a different type of modal vibrations in loudspeaker cones. The resonant modes of hoop stress induced vibration are referred to as being bell or radial modes.

Radial modes are particularly troublesome in that they occur at lower frequencies and require smaller amounts of energy to excite them as compared with the concentric modes. After an extensive study, Krüger has evolved a voice coil former, cone, and surround structure that is almost completely free of radial modes. Taken together, the resonant modes lead to what is referred to as coloration as the loudspeaker will have peaks and dips in the pressure response as a function of the driving frequency. Fig. 18-11 illustrates in a plan view of a loudspeaker cone the lowest order of both the concentric and radial resonance modes. The shading in the figure represents motion away from the observer while the unshaded portion represents motion towards the observer.

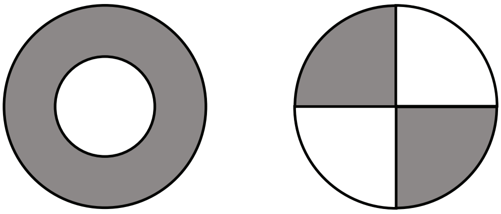

A. Concentric Mode B. Radial Mode

Figure 18-11. Lowest order resonance modes.

Dome Radiators

Many mid and high frequency devices do not employ the conical structure at all but rather have diaphragms that are shaped like domes or partial hemispheres. This is particularly true with regard to compression drivers employed with high frequency horns. Fig. 18-12 is a section view of such a device.

The driving force is applied by the voice coil at the base of the dome to initiate the acoustical disturbance in the material constituting the dome. This disturbance is communicated to the other parts of the dome with the speed of sound in the dome material. As long as the material has finite stiffness, this speed also will be finite and just as was the case with the cone, there is a transmission delay or phase lag between the various parts of the dome.

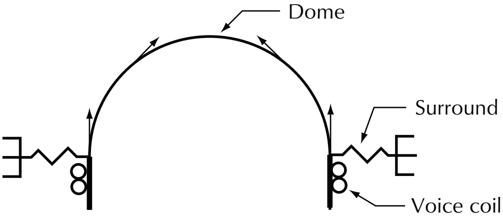


Figure 18-12. Dome radiator.

For an impulsive type of driving force, the disturbance starts out as a wrinkle around the base of the dome and moves across the dome surface in the direction indicated by the arrows. The circumference of the wrinkle is decreasing as it moves upward over the dome's surface, collapses on itself as it crosses over the dome's top, and expands again as it continues across to the other side. This implies that a disturbance originating on the extreme right in the figure must be terminated on the extreme left and vice versa. This means that the surround, which is attached at the base where the disturbance originates, must also act as the appropriate absorber for the disturbance once it has traversed the entire dome's surface. Again, just as in the case of the cone shape, there exists a termination problem. An improper termination results in wave reflection and introduces the possibility of resonant standing waves on the dome's surface. The behavior of a dome radiator is thus qualitatively similar to that of a cone differing only in the degree of participation of the phase lag and breakup effects. Dome style direct radiators excel in high frequency applications where their small size contributes to good angular dispersion.

Modern laser optic based instrumentation now allows detailed observation of diaphragm motion regardless of diaphragm shape and thus provides guidance in the search for improved diaphragm structures possessing minimal breakup effects. Phase lag effects will always remain, however, because of finite wave speed in the cone material.

18.2 Radiated Power

It is a relatively simple task to calculate the power radiated by a piston type device under sinusoidal excitation provided that both the piston velocity and the radiation resistance are known quantities. The relationships are

$$
\begin{aligned}
P_r &= {u_{rms}}^2 R_r \\
&= \frac{1}{2}{u_m}^2 R_r \\
&= \frac{1}{2}|u|^2 R_r
\end{aligned}
\tag{18-44}
$$

where,

P_r is the average radiated power,
u_{rms} is the root mean square velocity,
u_m is the velocity amplitude,
u is the piston velocity expressed as a phasor,
R_r is the radiation resistance.

The notation $|u|^2$ reads as the square of the absolute magnitude of the phasor representing the piston velocity. The value of the radiation resistance at any frequency may be calculated with the aid of Fig. 18-2. The determination of the piston velocity, however, is a more difficult problem. In solving this problem one must return to the basics and inquire what does a voltage source "see" when driving a piston-like loudspeaker. This is presented in Fig. 18-13.

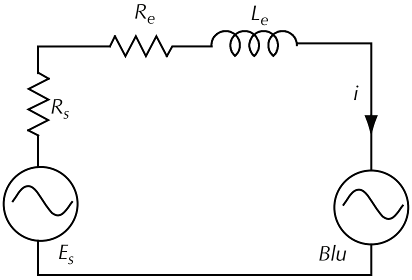


Figure 18-13. The circuit as "seen" by a voltage source driving a dynamic loudspeaker.

In the figure, E_s is the open circuit voltage or emf of the source, R_s is the source resistance, R_e is the electrical resistance of the voice coil, L_e is the electrical self-inductance of the voice coil, and Blu is the emf induced in the voice coil as a result of its motion in the magnetic field of the magnetic gap. As is true of any induced emf, Blu acts counter to the causative agent which is the current in the voice coil. As a result of this the expression for the voice coil current may be written as

$$
i = \frac{E_s - Blu}{R_s + R_e + j\omega L_e}
\tag{18-45}
$$

The mechanical driving force acting on the voice coil because of the current is just Bli and the velocity in the steady state that results from the application of this force is just the force divided by the mechanical impedance hence the velocity is given by

$$
u = \frac{Bli}{Z_m}
\tag{18-46}
$$

The mechanical impedance at frequencies below and immediately above driver resonance is adequately represented by Eq. 18-14. This will need some adjustment at higher frequencies and will be so noted where appropriate. When Eqs. 18-45 and 18-46 are solved simultaneously for the piston velocity, the resulting expression is found to be

$$
u = \frac{\dfrac{E_s Bl}{R_s + R_e + j\omega L_e}}{\dfrac{B^2 l^2}{R_s + R_e + j\omega L_e} + R_m + R_r + j\left[\omega M + X_r - \dfrac{K + K_b}{\omega}\right]}
\tag{18-47}
$$

The numerator of Eq. 18-47 has the dimensions of mechanical force or Newtons while the denominator has the dimensions of mechanical impedance or kg/s. The quotient of the two has the dimensions of mechanical velocity or m/s. It is possible to construct an analogous "mechanical circuit" which may be analyzed to obtain the result of Eq. 18-47. Such a circuit appears as Fig. 18-14.

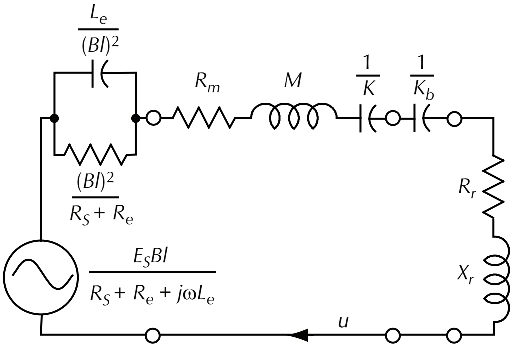


Figure 18-14. Loudspeaker analogous circuit model.

In a purely electrical case, one determines the circuit current by dividing the applied emf by the total circuit electrical impedance. Similarly, one determines the mechanical velocity in this model by dividing the mechanical force of the generator by the total mechanical impedance of the "circuit." It is instructive to see how this model conforms to the conditions stated in Eq. 18-47. Obviously, the force term is the term written adjacent to the generator in

Fig. 18-14. How then does one arrive at an expression for the total mechanical impedance? The answer lies in the choice of symbols in the model. The symbol for a resistor is manipulated as one would do for a resistor in an electrical circuit and similarly for capacitors and inductance. For example, consider the parallel combination of resistance and capacitance immediately to the right of the generator in the figure. The rule is

$$
\begin{aligned}
\frac{1}{Z_t} &= \frac{1}{Z_R} + \frac{1}{Z_C} \\
&= \frac{1}{\dfrac{B^2 l^2}{R_e + R_s}} + \frac{1}{\dfrac{-jB^2 l^2}{\omega L_e}} \\
&= \frac{R_e + R_s + j\omega L_e}{B^2 l^2}
\end{aligned}
$$

One needs only to take the reciprocal of the final term above to obtain

$$Z_t = \frac{B^2 l^2}{R_e + R_s + j\omega L_e}$$

This result is the first term appearing in the denominator of Eq. 18-47. The remaining terms in the circuit are series connected and treated as such. Upon combining the resistive elements one has

$$R_r + R_m$$

The positive reactance terms are inductors and give a total

$$j\omega M + jX_r$$

The capacitive terms combine as capacitors connected in series giving

$$
\begin{aligned}
\frac{1}{C_t} &= \frac{1}{\dfrac{1}{K}} + \frac{1}{\dfrac{1}{K_b}} \\
&= K + K_b \\
&= K'
\end{aligned}
$$

Thus the total capacitance is the reciprocal of the last term above or

$$\frac{1}{K'}$$

The impedance of such a capacitance is

$$
\begin{aligned}
Z_C &= \frac{-j}{\omega C} \\
&= \frac{-jK'}{\omega}
\end{aligned}
$$

The total mechanical impedance of the model is just the sum of all of the individual terms found above. The addition of these terms yields

$$\frac{B^2 l^2}{R_s + R_e + j\omega L_e} + R_m + R_r + j\left[\omega M + X_r - \frac{K + K_b}{\omega}\right]$$

This final expression is just the term in the denominator of Eq. 18-47. The foregoing analysis illustrates the validity of the model.

The employment of the full model as displayed in Fig. 18-14 is not always required in order to obtain useful information. One almost universal simplification is associated with the appearance of the source resistance R_s.

Most modern power amplifiers have source resistances that are only a small fraction of an ohm and hence are considerably smaller than typical voice coil resistances that are several ohms in size. If the wiring between the amplifier and the loudspeaker also represents a fraction of an ohm, then the omission of the source resistance creates negligible error.

Additionally, depending upon the frequency range of interest, other simplifications are possible. For example, in the low frequency range well below driver resonance where ω is quite small, the term ωL_e is also quite small and can be ignored. Furthermore, the mechanical impedance represented by the series connected capacitances is considerably larger than that of the other elements. For the purposes of calculating the piston velocity in this range the greatly simplified model of Fig. 18-15 can be employed.

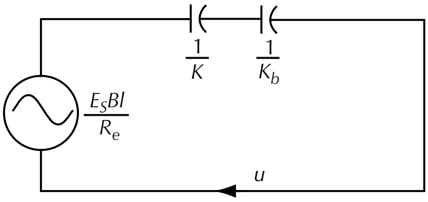


Figure 18-15. Loudspeaker model at very low frequencies.

Employing Fig. 18-15, the velocity at very low frequencies is given by,

$$u = \frac{\frac{E_sBl}{R_e}}{\frac{-jK'}{\omega}} = \frac{jE_sBl}{R_eK'}\omega \tag{18-48}$$

where,
u is the phasor representing the piston velocity,
E_s is the phasor representing the open circuit voltage of the electrical source driving the loudspeaker.

This equation as it stands gives both amplitude and phase information. For the purposes of a power calculation one requires only the absolute magnitude or amplitude of the piston velocity expressed as

$$u_m = \frac{E_{sm}Bl}{R_eK'}\omega \tag{18-49}$$

where,
E_{sm} is the open circuit voltage amplitude of the generator.

The conclusion here is that the velocity amplitude is directly proportional to the angular frequency. The radiation resistance in this frequency range is found from

$$R_r = \frac{\rho_0S^2\omega^2}{2\pi c} \tag{18-50}$$

According to Eq. 18-44, then, the average radiated power in this frequency range will be

$$P_r = \left(\frac{E_{sm}Bl}{R_eK'}\right)^2\frac{\rho_0S^2}{4\pi c}\omega^4 \tag{18-51}$$

If one increases the operating frequency in this range by a factor of two or a span of an octave, the radiated power will increase by a factor of two raised to the fourth power or 16. This is a power increase of 12 dB over a span of one octave hence, the power curve rises at the rate of 12 dB/octave in the very low frequency range.

In the range of frequencies immediately below and immediately above the driver resonant frequency the significant terms of the model are different. Here the voice coil inductance and the radiation resistance can still be neglected in calculating the velocity. The model now becomes that of Fig. 18-16.

In making calculations of the velocity in this frequency range, the radiation reactance X_r combines with M to become M' and the two capacitors combine to form $1/K'$. The velocity is then found from

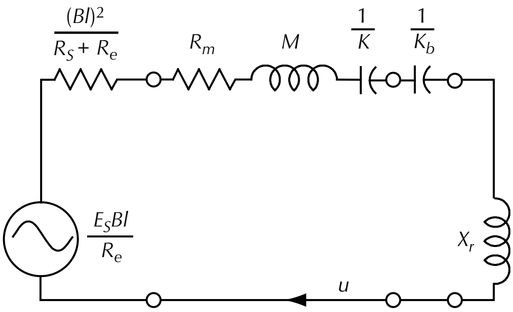


Figure 18-16. Model for frequencies in the vicinity of driver resonance.

$$u = \frac{\frac{E_sBl}{R_e}}{R_m + \frac{B^2l^2}{R_e} + j\left[\omega M' - \frac{K'}{\omega}\right]} \tag{18-52}$$

As discussed earlier, the velocity in this region may exhibit a peak, flatten out smoothly and rapidly, or slowly approach a nearly constant value depending upon the system total quality factor as in the earlier discussion. At resonance, of course, the reactance becomes zero and the average radiated power at resonance becomes

$$P_r = \left(\frac{\frac{E_{sm}Bl}{R_e}}{R_m + \frac{B^2l^2}{R_e}}\right)^2\frac{\rho_0S^2}{4\pi c}\omega_o^2 = \left(\frac{E_{sm}Bl}{R_eM'}\right)^2\frac{\rho_0S^2}{4\pi c}Q_t^2 \tag{18-53}$$

where,
Q_t is the driver total quality factor.

In the intermediate frequency range above resonance, the significant components of the model are again different. A significant component is one whose impedance is large enough that it must be considered in making numerical calculations. The model in this frequency range has the appearance presented in Fig. 18-17.

In this frequency range, the impedance of the mass and radiation reactance are still directly proportional to the frequency and are much larger than any of the other circuit impedances. The radiation resistance is still growing as the frequency squared but is small. The other impedance terms are constants as is the force of the generator. The net

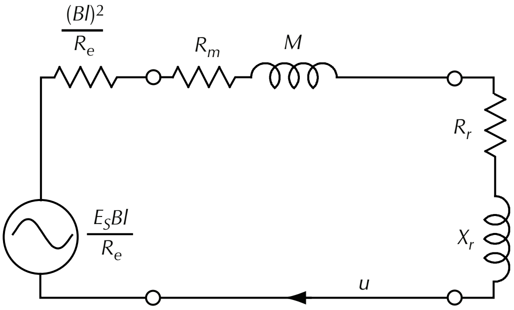


Figure 18-17. Significant model features for intermediate frequencies above resonance.

result is that u_m is proportional to reciprocal frequency and u_m^2 is then proportional to the square of the reciprocal frequency. Therefore,

$$u_m^2 \propto \frac{1}{\omega^2}$$

$$R_r \propto \omega^2 \tag{18-54}$$

$$P_r = \frac{1}{2}R_r u_m^2 \propto \text{a constant}$$

In this frequency range, then, the radiated power is a constant. Also in this range, one first encounters the effects of radial and concentric resonant modes of loudspeaker cones. Break up resonant modes are not explained by the present model as it is fashioned after a true rigid piston radiator.

As the high frequency range is approached the significant components of the model undergo another shift as displayed in Fig. 18-18.

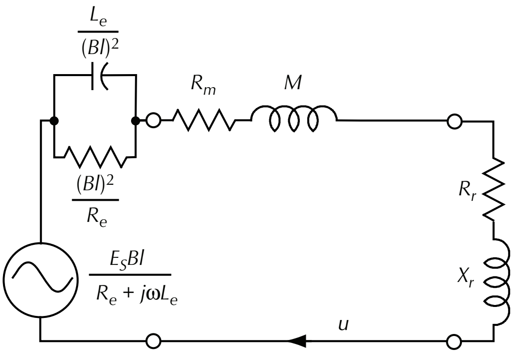


Figure 18-18. Significant model features in the near high frequency range.

In this range the effect of the voice coil self-inductance becomes of importance. In fact, if the voice coil inductance is large enough such that the capacitance in the parallel combination is dominant over the parallel resistance, there exists the possibility of another resonance between the negative reactance of this capacitance with the positive reactance of the mass inductance and the radiation reactance. Also in this range, the radiation resistance is ceasing its growth and is approaching a constant value while the radiation reactance is diminishing. The resonance, if it occurs, will give a lift and a peak to the power curve in this region. On the other hand, if the parallel resistance term dominates over the capacitive term, this second resonant peak will not appear and the radiated power will begin to slowly diminish.

Finally in the high frequency range the inductive reactance of the force generator becomes much larger than the voice coil resistance. The impedance of the capacitance in the parallel combination becomes quite small and thus approaches being a short circuit as compared with the parallel resistance. The positive reactance of the mass being directly proportional to frequency becomes very large and dominates over other impedance terms. The radiation reactance has become quite small in this frequency range and the radiation resistance has attained a constant value of $\rho_0 cS$. Additionally, the wavelengths are now smaller than the dimensions of the back enclosure and the piston is now radiating into the back cavity as well as from its exposed face so that the model radiation resistance and reactance must now be doubled. Provision must be made in the back cavity to absorb this radiant energy through the employment of a dissipative lining. The useful acoustic power remains that which is radiated from the exposed face of the piston. This form of the model is presented in Fig. 18-19.

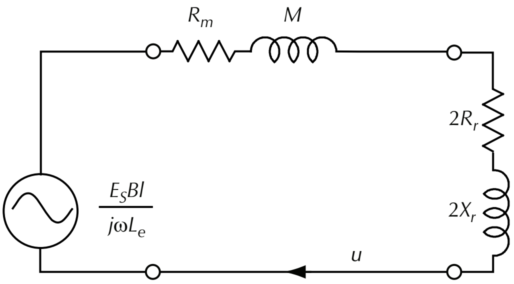


Figure 18-19. Model in the high frequency range.

An examination of Fig. 18-19 indicates that the driving force is inversely proportional to the frequency. The circuit impedance is dominated by the mass reactance and this term grows linearly with frequency. The piston velocity depends on the quotient of the force and the impedance and thus is inversely proportional to the square of the frequency. The radiation resistance is now a constant. The radiated power is proportional to the product of the radiation resistance and the square of the velocity with the result that the radiated power is decreasing with the fourth power of the frequency. This is a rate of –12 dB/octave.

18.3 Axial Sound Pressure Level

The axial sound pressure level in the far field in all instances is proportional to the product ωu_m. In the preceding analysis it was found that in the range well below driver mechanical resonance that the velocity was growing linearly with frequency which means that the pressure is increasing as the square of the frequency. This implies that the level is growing at the rate of 12 dB/octave as was true also of the power. The sound pressure level curve in fact mimics that of the power level curve everywhere except in the high frequency range. In the high frequency range the velocity amplitude is falling as the square of the frequency. The pressure, however, depends upon the product of the velocity amplitude with the frequency. This product is inversely proportional to frequency and thus has a rate of only –6 dB/octave.

Many attribute this departure between the power curve and the pressure curve to the increasing directivity of the radiator at high frequencies. This departure, as shown, depends on the behavior of the velocity rather than increasing directivity as the directivity function on axis has a value of unity at all frequencies. The behavior of both the radiated power and the axial sound pressure for a driver optimized for mid bass operation is displayed in Fig. 18-20.

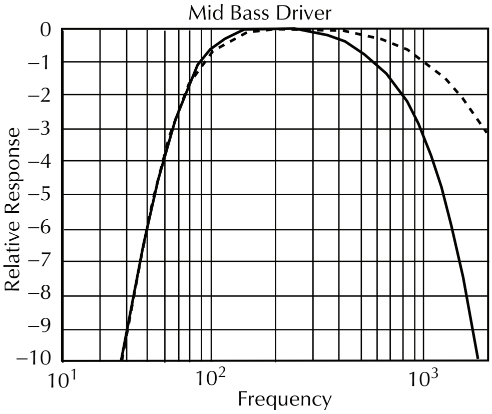


Figure 18-20. Mid bass driver calculated power and pressure responses.

18.4 Efficiency

A loudspeaker efficiency coefficient is the ratio of the average acoustical power radiated by the loudspeaker to the total average power supplied to the loudspeaker by the driving source. If the average radiated power is denoted as $<P_r>$ and the average electrical power supplied by the source is denoted as $<P_e>$ then the efficiency coefficient η is simply

$$\eta = \frac{<P_r>}{<P_e>} \tag{18-55}$$

The efficiency when expressed as a percentage is this number multiplied by 100. A portion of the electrical power supplied is wasted as heat in the voice coil electrical resistance. A further portion appears as heat in the mechanical resistance of the suspension. Lastly, a portion of the electrical power supplied is converted into useful radiated power. It is possible to express the efficiency coefficient in terms of the loudspeaker parameters in the following way. First it is necessary to relate the piston velocity to the voice coil current. The phasor representing the piston velocity is given by the quotient of the driving force with the mechanical impedance of the loudspeaker.

$$u = \frac{Bli}{Z_m} \tag{18-56}$$

The term in the denominator of Eq. 18-56 is the mechanical impedance first presented in Eq. 18-14. In making power calculations one requires only the velocity amplitude which is

$$u_m = \frac{Bli_m}{|Z_m|} \tag{18-57}$$

The denominator of Eq. 18-57 is the magnitude of the complex mechanical impedance. The average radiated power is now

$$<P_r> = \frac{1}{2}R_r u_m^2 \tag{18-58}$$

If one were to employ the rms value of the velocity rather than the amplitude, the factor of 1⁄2 would not appear in Eq. 18-58. The average electrical power as stated earlier consists of the sum of three terms. The first of these terms is the rate of heat production in the voice coil resistance

$$<P_1> = \frac{1}{2}i_m^2 R_e \tag{18-59}$$

The second of these terms is the rate of heat production in the mechanical resistance of the suspension

$$\langle P_2 \rangle = \frac{1}{2} u_m{}^2 R_m \tag{18-60}$$

The final term is the average radiated power of Eq. 18-58. Thus far then, the efficiency coefficient can be written as

$$\eta = \frac{u_m{}^2 R_r}{i_m{}^2 R_e + u_m{}^2 R_m + u_m{}^2 R_r} \tag{18-61}$$

where the common factor of ½ has been dropped in writing Eq. 18-61.

Finally, upon solving Eq. 18-57 for the current amplitude and substituting into Eq. 18-61, the final result appears as

$$\eta = \frac{R_r}{\frac{|Z_m|^2}{(Bl)^2} R_e + R_m + R_r} \tag{18-62}$$

The efficiency coefficient is highly frequency dependent and reaches its maximum value in the vicinity of the mechanical resonance of the piston. The plot of the efficiency coefficient curve of the mid-bass loudspeaker of Fig. 18-20 appears in Fig. 18-21.

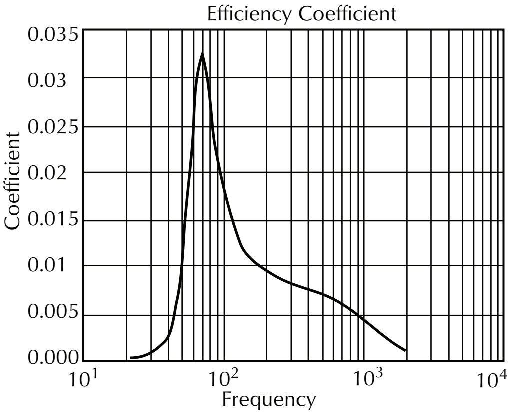


Figure 18-21. Mid bass efficiency coefficient curve.

The maximum efficiency of this particular driver is just a little over 3%. This low value is not unusual for direct radiator loudspeakers.

18.5 Loudspeaker Electrical Impedance

The electrical impedance of a loudspeaker when the cone is clamped preventing cone motion is simply $R_e + j\omega L_e$. When the cone is free to move, however, the counter emf induced by the voice coil's motion in the magnetic field greatly alters the impedance as viewed by a voltage source driving the loudspeaker. The electrical impedance presented to the voltage source under these conditions is now

$$Z_e = R_e + j\omega L_e + \frac{B^2 l^2}{Z_m} \tag{18-63}$$

This electrical impedance behavior is usually displayed in the form of a plot of the magnitude of this complex quantity versus the frequency or log frequency. Such a theoretical plot for the mid bass driver of Fig. 18-20 appears in Fig. 18-22.

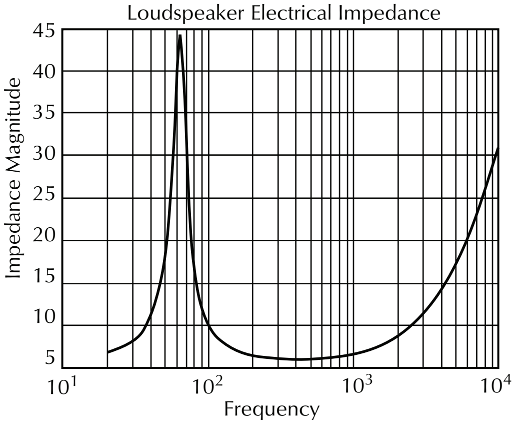


Figure 18-22. Loudspeaker electrical impedance.

The maximum on the curve occurs essentially at the mechanical resonant frequency and is significantly larger than the minimum of the curve that occurs above the mechanical resonance. The minimum on the curve is usually taken to be what is called the nominal impedance of the loudspeaker. The rising impedance at high frequencies is contributed by the voice coil inductance. If this had been a measured rather than theoretical curve, the presence of break up modes would be evident by a series of small bumps beginning in the vicinity of 400 Hz and continuing into the higher frequency regions. This topic is discussed in greater detail in Chapter 8 *Interfacing Electrical and Acoustic Systems*.

18.6 Loudspeaker Directivity Factor

Acoustic intensity is a vector quantity that describes both the magnitude and direction of the instantaneous acoustic energy flow per unit area per unit time. Its direction at any point is the direction of acoustic wave propagation at that point. The time average of the magnitude of the intensity, which is not a vector quantity, can readily be related to the acoustic pressure at the point in question. In the case of sinusoidal excitation this relation between the average intensity and the acoustic pressure is

$$\langle I \rangle = \frac{1}{2}\frac{p_m^2}{\rho_0 c} \tag{18-64}$$

where,

p_m is the amplitude of the sinusoidal acoustic pressure.

As usual, if one employs the rms acoustic pressure, the factor of 1/2 can be dropped from the equation. The concept of source directivity can be applied to any acoustic source regardless of its shape. The directivity in acoustic work, unfortunately, is usually symbolized by the letter Q. This is unfortunate in the sense that Q with various subscripts is also used to denote quality factor which is an entirely different concept. In words, Q is the ratio of the average intensity at some point in the far field of the actual source divided by the average intensity at the same point that would be produced by a point source radiating the same total power as does the actual source. Point sources of course radiate isotropically. This means the radiation is uniform in all directions. Q, then, is written as

$$Q(\theta, \varphi) = \frac{\frac{1}{2} \times \frac{p_m^2(r, \theta, \varphi)}{\rho_0 c}}{\frac{\langle P \rangle}{4\pi r^2}} \tag{18-65}$$

The variables r, θ, and φ define the location of the point in space where Q is being evaluated relative to a spherical polar coordinate system fixed in the source. The symbol $\langle P \rangle$ represents the average radiated power. Even though not indicated in the equation, Q is also a function of the frequency as both acoustic power and acoustic pressure are frequency dependent. The term in the denominator of Eq. 18-65 is the average intensity of a point source. Both the numerator and denominator of Eq. 18-65 vary inversely with r^2 so Q is independent of r provided that the evaluation is made in the far field. If the source is symmetrical about the polar axis as would be the case for a direct radiator then Q will depend on the polar angle θ but will be independent of the azimuthal angle φ. The axial Q of such a radiator is the value of Q when θ is set equal to zero. Fig. 18-23 displays the behavior of the axial Q of a piston radiating into a half space.

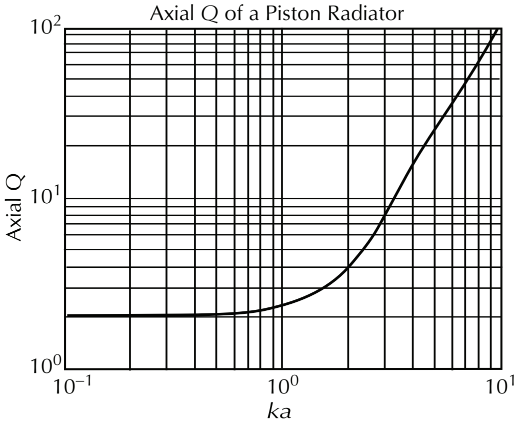


Figure 18-23. Axial Q of a piston radiator.

In viewing Fig. 23 it should be remembered that $k = \omega/c$ and thus the values of the abscissa are proportional to the product of the operating frequency with piston radius and thus are applicable to all piston sizes. For small values of the abscissa Q has a nearly constant value of 2. Q begins to grow when the abscissa exceeds one and becomes quite large as the abscissa approaches ten or greater.

An often used quantity related to the axial Q is called the Directivity Index symbolized as *DI*. The definition of *DI* is given by Eq. 18-66.

$$DI = 10 \text{ dBlog } Q \tag{18-66}$$

18.7 Loudspeaker Sensitivity

A specification of loudspeaker sensitivity is a way of stating what axial sound pressure level will be produced at a given distance by some standard electrical input to the loudspeaker. The distance most usually specified is 1 m but that does not mean that the measurement is made at 1 m. In order to properly assess the loudspeaker performance the measurement must be made in the far field. If r is the far field distance expressed in meters, then the pressure level at 1 m is the pressure level at a distance r plus the correction which accounts for the inverse variation of pressure with distance so

$$L_p\ @1\text{ m} = L_p(r) + 20 \text{ dBlog}\left(\frac{r}{1\text{ m}}\right) \tag{18-67}$$

The standard electrical input is most often stated as being an average power of 1 W although the common practice has become one where the power supplied is not a true measured watt. What is usually applied to the loudspeaker is a specified band of pink noise. The rms voltage of this pink noise is adjusted so that it would produce an average power of 1 W in a pure resistance equal to the rated impedance of the loudspeaker. This rated impedance is somewhat at the discretion of the manufacturer and may or may not be the nominal impedance of the loudspeaker as extracted from a measured impedance curve. If the average power as determined above is different from 1 W the pressure level must be adjusted according to 10 dBlog(<P>/1 W).

Therefore if L_p is the sensitivity of the loudspeaker at 1 m for 1 W, then the far field sound pressure level will be given by

$$L_p(r) = L_p - 20\text{ dBlog}\left(\frac{r}{1\text{ m}}\right) + 10\text{ dBlog}\left(\frac{<P>}{1\text{ W}}\right) \tag{18-68}$$

18.8 Direct Radiator Example Calculations

The loudspeaker data employed in the following calculations unless otherwise pointed out are extracted from Table 18-1. Additionally, static air pressure will be taken as 1.013×10^5 pascals, static air density as 1.2 kg/m^3, and sound speed as 344 m/s.

1. What air volume in the back enclosure employed with this driver will allow maximally flat response for half space radiation at low frequencies?

   Solution: The governing equation is Eq. 18-32 which contains the unknown term Q_{fa} which in turn is given by Eq. 18-28. Eq. 18-28 contains the unknown term M' which in turn is given by Eq. 18-17. Eq. 18-17 then serves as the starting point.

$$\begin{aligned} M' &= 0.03 + \frac{(1.2)(8)(\pi)(0.1)^2(0.1)}{3\pi} \\ &= 0.0332 \end{aligned}$$

$$\begin{aligned} Q_{fa} &= \frac{\sqrt{(3000)(0.0332)}}{2.5 + \dfrac{(1^2)(10^2)}{6}} \\ &= 0.521 \end{aligned}$$

$$\begin{aligned} V_0 &= \frac{(1.4)(1.013 \times 10^5)((\pi)(0.1^2))^2}{3000\left[\left(\dfrac{0.7071}{0.521}\right)^2 - 1\right]} \\ &= 0.0554\text{ m}^3 \end{aligned}$$

2. What is the frequency of mechanical resonance of this driver when mounted in the back enclosure that yields maximally flat response?

   Solution: The governing equation here is Eq. 18-21 that contains the unknown term K' which is given by Eq. 18-4. Eq. 18-4 contains the unknown term K_b that in turn leads to Eq. 18-3. Eq. 18-3 then serves as the starting point.

$$\begin{aligned} K_b &= \frac{(1.4)(1.013 \times 10^5)((\pi)(0.1^2))^2}{0.0554} \\ &= 2527\text{ N/m} \end{aligned}$$

$$\begin{aligned} K' &= 3000 + 2527 \\ &= 5527 \end{aligned}$$

$$\begin{aligned} f_0 &= \frac{1}{2\pi}\sqrt{\frac{5527}{0.0332}} \\ &= 65\text{ Hz} \end{aligned}$$

3. Assuming that the driver is appropriately back enclosed, what is the radiation resistance for half space radiation at the frequency of mechanical resonance?

   Solution: The governing equation is Eq. 18-50 with ω set equal to ω_0.

$$\begin{aligned} R_r &= \frac{1.2((\pi)(0.1)^2)^2(2\pi 65)^2}{2\pi 344} \\ &= 0.0914\text{ kg/s} \end{aligned}$$

4. 4. If a 20 V_{rms} sinusoid is applied to this loudspeaker at the frequency of mechanical resonance, what will be the axial sound pressure level at a distance of 5 m?

   Solution: One can employ Eq. 18-24 with the help of Eq. 18-25 to determine either the pressure amplitude or rms pressure. As the voltage is stated as being rms, its employment will lead to an rms value of the pressure. Evaluation of the first bracketed term in Eq. 18-24 yields

$$\frac{((20)(1)(10)(1.2)(\pi)(0.1)^2)}{2\pi(5)(0.0332)(6)} = 1.2\text{ Pa}$$

Upon setting $\omega = \omega_0$, the magnitude of the second bracketed term in Eq. 18-24 becomes

$$\frac{1}{\sqrt{(1-1)^2 + \frac{1}{(0.7071)^2}}} = 0.7071$$

The rms pressure is then the product of these two or 0.849 pascal. The sound pressure level is then

$$\begin{aligned} L_p &= 20\log\frac{0.849}{2\times 10^{-5}} \\ &= 93\text{ dB} \end{aligned}$$

5. What is the efficiency coefficient of this loudspeaker at the frequency of mechanical resonance?

Solution: From Eq. 18-16 the mechanical impedance at resonance is purely real and is just the sum of the mechanical resistance and the radiation resistance. The radiation resistance was calculated in problem 3. Therefore making use of Eq. 18-62

$$\begin{aligned} \eta &= \frac{0.0914}{\frac{(2.5+0.0914)^2}{10^2}(6) + 2.5 + 0.0914} \\ &= 0.0305 \end{aligned}$$

6. What is the average power radiated by this loudspeaker when driven by the 20 V sinusoid at its frequency of mechanical resonance?

Solution: The axial sound pressure at 5 m found in problem 4 is 0.849 Pa. The axial intensity at 5 m is then

$$\begin{aligned} I &= \frac{p^2}{\rho_0 c} \\ &= \frac{(0.849)^2}{(1.2)(344)} \\ &= 1.75\times 10^{-3}\text{ W/m}^2 \end{aligned}$$

The radiation at this low frequency is isotropic in a half space hence the average radiated power is the intensity multiplied by the area of a hemisphere whose radius is 5 m or

$$\begin{aligned} P &= I(2\pi r^2) \\ &= (1.75\times 10^{-3})(2\pi 5^2) \\ &= 0.275\text{ W} \end{aligned}$$

7. What average electrical power must be supplied to produce the radiated acoustical power of problem 6?

Solution: The input electrical power is the radiated power divided by the efficiency coefficient.

$$\begin{aligned} P_e &= \frac{P_r}{\eta} \\ &= \frac{0.275}{0.0305} \\ &= 9.02\text{ W} \end{aligned}$$

8. What are the piston velocity amplitude and the piston displacement amplitude under the conditions of problem 6?

Solution: The pressure amplitude is the rms pressure multiplied by $\sqrt{2}$.

$$\begin{aligned} p_m &= p_{rms}\sqrt{2} \\ &= 0.849\sqrt{2} \\ &= 1.2\text{ Pa} \end{aligned}$$

$$\begin{aligned} u_m &= \frac{p_m 2\pi r}{\rho_0 \omega S} \\ &= \frac{1.2(2\pi)(5)}{1.2(2\pi 65)(\pi)(0.1)^2} \\ &= 2.45\text{ m/s} \end{aligned}$$

The displacement amplitude is the velocity amplitude divided by the angular frequency.

$$\begin{aligned} \xi_m &= \frac{u_m}{\omega} \\ &= \frac{2.45}{2\pi(65)} \\ &= 6\times 10^{-3}\text{ m} \end{aligned}$$

9. The subject back enclosed loudspeaker is now suspended in space far from any external bounding surface such that it now radiates into all of space. What changes now occur in the operation of the loudspeaker assuming the applied voltage remains the same?

Solution: The starting point here is the radiation impedance. In going from half space to all space the radiation impedance will be halved. This means that M' will be a slightly different value because Eq. 18-17 now becomes

$$\begin{aligned} M' &= M + \frac{\rho_0 4Sa}{3\pi} \\ &= 0.03 + \frac{1.2(4)(\pi)(0.1)^2(0.1)}{3\pi} \\ &= 0.0316\ \text{kg} \end{aligned}$$

This causes a modest change in the frequency of resonance.

$$\begin{aligned} f_0 &= \frac{1}{2\pi}\sqrt{\frac{K'}{M'}} \\ &= \frac{1}{2\pi}\sqrt{\frac{5527}{0.0316}} \\ &= 66.6\ \text{Hz} \end{aligned}$$

The driving signal is still at 65 Hz and hence the loudspeaker is operating slightly below its present resonant frequency. The change in the overall mechanical impedance, however, is so small that the piston velocity and displacement are changed by only a negligible amount. Even though the piston velocity is essentially unchanged the acoustic pressure is reduced by a factor of two because the radiation now is into a whole space rather than a half space. The radiation resistance is now half of its former value. As the piston velocity is essentially unchanged, the radiated power will track the change in radiation resistance and will become half of its former value. The efficiency coefficient is also essentially half of its former value. As a consequence, the required electrical power is changed by a negligible amount. The fact that the axial acoustic pressure is halved means that the sound pressure level is reduced by 6 dB and thus becomes 87 dB.

10. The back enclosure is now removed from this driver and the driver is mounted in the center of a large flat wall that serves as a partition between two large rooms. Alternatively, the driver might be mounted in a rigid ceiling having a large air space above it. What are the operational conditions of the driver under these mounting conditions?

Solution: The driver now radiates freely into a half space on both its front and rear faces. The radiation impedance acting on each face is given by Eq. 18-5 and hence the total radiation impedance acting on the driver is just twice that given by Eq. 18-5. At low frequencies the mass contributed by the radiation reactance is twice the former value so that M' becomes

$$\begin{aligned} M' &= M + \frac{2\rho_0 8Sa}{3\pi} \\ &= 0.03 + 0.0064 \\ &= 0.0364 \end{aligned}$$

Additionally, there is now no stiffness contributed by a back enclosure so that K' is identical to K. The frequency of mechanical resonance will now be

$$\begin{aligned} f_0 &= \frac{1}{2\pi}\sqrt{\frac{K}{M'}} \\ &= \frac{1}{2\pi}\sqrt{\frac{3000}{0.0364}} \\ &= 45.7\ \text{Hz} \end{aligned}$$

The total quality factor as calculated by Eq. 18-22 has a new value.

$$\begin{aligned} Q_t &= \frac{\omega_0 M}{R_m + \dfrac{B^2 l^2}{R_e}} \\ &= \frac{2\pi(45.7)(0.0364)}{2.5 + \dfrac{100}{6}} \\ &= 0.545 \end{aligned}$$

The new total quality factor is less than the optimum value of 0.7071 so that the low end response is no longer maximally flat. It is possible to raise the quality factor by connecting additional resistance in series with the voice coil thus making R_e appear to be larger. If a resistor of 2.14 Ω is added in series with the loudspeaker thus making R_e appear to be 8.14 Ω, the total quality factor will become the optimum value for maximally flat response. The adverse effect of raising the total quality factor in this fashion is the significant reduction that occurs in the efficiency coefficient.

18.9 Horns and Compression Drivers

Horns play two fundamental roles in acoustics. Firstly, they are directional control devices serving to guide the airborne acoustic energy into particular directions or regions and as such they might well be called waveguides. The familiar plane wave tube may well be considered to be a special case of a horn as it certainly constitutes a waveguide. Secondly, horns act as impedance matching devices similar to the action of an electrical transformer.

Horn surfaces define a bounded region whose cross-sectional area increases from the input to the output in a loudspeaker application. The acoustic power flowing through a cross section of area *S*, when the acoustic pressure and particle velocity are in phase, is the product puS. The product uS is called the volume velocity and is denoted by U therefore the acoustic power flow is pU. At the input end of the horn where S is small and the acoustic pressure p is large, the volume velocity for a given acoustic power is small.

At the output end of the horn where S is large, the volume velocity is large and the acoustic pressure is small for the same acoustic power. This behavior is analogous to an electrical step down transformer that has a large voltage and small current in the primary and a small voltage and large current in the secondary.

Horns were used in acoustics long before their principles of operation were even partially understood. The earliest hearing aid devices were based on horns with sound transmission in the reversed sense in that the object was to convert large U and small p into large p with small U to accommodate the fact that the ear is a pressure sensitive organ. Some of the earliest acoustic recordings also employed horns operating in reverse with the mouth of the horn collecting energy over a large area and concentrating it into a small area to actuate a small diaphragm mechanically coupled to a recording stylus. This procedure was reversed in the reproduction process wherein the horn was employed in the more conventional sense.

Horns have also been employed as the basis of many musical instruments. The requirements of a horn to be used for sound generation are radically different from that of sound reproduction or reinforcement. In the case of sound generation, resonances in the horn are desirable and in fact, essential. In sound reproduction or reinforcement, however, resonances are undesirable and steps must be taken to minimize their existence. This underscores the necessity for having an underlying theory of horn operation to guide the construction for various applications.

Horn theory stems from the original work of Euler, Lord Rayleigh, and Webster. Webster was the first to introduce the concepts of specific acoustic impedance and analogous acoustic impedance, both of which have proven to be very valuable in acoustic analysis. What is known as Webster's horn equation is a wave equation that in the strictest sense is correctly applicable to only three waveguide structures. These are the plane wave tube, the conical horn, and the cylindrical horn. Webster's equation employs only a single space variable in the axial direction implying that the acoustic pressure is uniform over an appropriately drawn cross section of the guide or horn structure. This is satisfied exactly in a plane wave tube of limited diameter that is excited at the input by a plane wave.

The equation is also exact for a conical horn that is excited with a spherical wave and for a cylindrical horn that is excited by a cylindrical wave. The application of the equation to other horn structures is only approximately correct and then only when the horn opens up or flares very slowly. It is this last requirement that is often lost sight of in practice.

The simplest horn geometry is that of a truncated cone wherein acoustic energy in the form of a diverging spherical wave is introduced into the small end of the cone and subsequently propagates freely within the cone as an outgoing wave as suggested in Fig. 18-24.

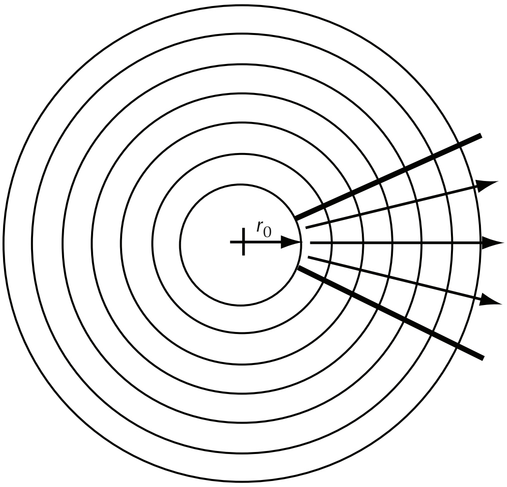


Figure 18-24. A spherical wave has a natural coordinate fit in a conical horn.

In order for this horn to work properly, the acoustic wave introduced at its small or throat end must have a radius of curvature equal to r_0, where r_0 is measured from the virtual apex of the truncated cone. The "natural fit" means that the spherical wavefronts diverging from the apex would everywhere be normal to the bounding surface provided by the cone and all energy flow would be radially directed. The particle velocity in such a wave motion would be tangent to the interior walls of the horn and there would be no reflections at the interior wall surface. Internal reflections occur when the particle velocity has a component normal to an interior wall surface. Plane waves, as another example, form a natural fit in a straight tube of constant cross section. In both instances the wave motion can be described using only a single spatial coordinate and as such are called single parameter waves. A third geometry that provides a natural fit is that of a horn formed from a sector of a cylinder. Here, however, the wave introduced at the throat must have a cylindrical wavefront. So far as is presently known these three geometries are the only ones that satisfy the natural fit or single parameter conditions exactly.

As a straight tube of constant cross section of course is not a horn in the strictest sense, that leaves only two natural horns. Many other horn shapes have been employed with varying degrees of success, however, but they are only approximately single parameter devices and all suffer from bounding wall reflections to a greater or lesser degree. An analysis by Morse concludes that in order for a horn shape to approximately fall under the conditions necessary to satisfy Webster's horn equation, the rate of change of the square root of the cross-sectional area with respect to the single axial parameter must be much less than one. To what degree some common horns satisfy or fail to satisfy this criterion is a point worthy of examination.

Webster's equations for single parameter horns are

$$\frac{1}{S}\times\frac{\partial\left(S\frac{\partial p}{\partial \chi}\right)}{\partial \chi} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} \qquad (18\text{-}69)$$

$$\frac{\partial u}{\partial t} = -\frac{1}{\rho_0}\frac{\partial p}{\partial \chi}$$

In these equations χ is the appropriate single space variable and the other symbols have their usual meanings. The area symbol S must be treated carefully. In a true one parameter horn it is the surface area of the appropriately shaped wavefront as a function of position. In other horns the assumption is made that the wavefronts are approximately plane and S is the true cross sectional area of the horn as a function of position.

The first of Eq. 18-69 when applied to the conical horn becomes just the spherical wave equation. The second of these equations allows the determination of the specific acoustic impedance at the throat of the horn. The specific acoustic impedance is the ratio of the complex acoustic pressure to the complex particle velocity. The mechanical impedance is thus the product of the specific acoustic impedance with the wavefront area S. With a suitable horn driver, these expressions are identical to those of a pulsating sphere of radius r_0 with the exception that all of the energy diverges through the horn rather than through the surface area of a sphere of radius r_0. The sound intensity on the axis of the horn is thus increased by the ratio of the area of the sphere to the wave entrance area of the horn. This ratio is $2/(1-\cos\theta)$ where θ is the angle between the horn axis and the interior surface of the horn.

For example, if the coverage angle of the horn is 40°, θ is 20° and this ratio is 33.16. This number is identical to the axial Q of the horn. This corresponds to a pressure level increase of about 15 dB on the axis of the horn as compared with that produced by a pulsating sphere without the aid of a horn. Thus far the horn has been treated as if there were only an outgoing wave. This would strictly be true if the horn were infinitely long. For any horn of finite length a reflection will occur at the mouth and a portion of the original energy will be directed along the horn back towards the driver. Such reflections can lead to the production of standing waves having resonant frequencies related to the horn length. It has been found in practice that there is only a negligible mouth reflection if the mouth perimeter is about 3 times the free space wavelength at the frequency of operation. Horns intended for use at low frequencies are thus large, unwieldy devices.

Salmon has described a family of horns that can have approximately single parameter behavior when they flare slowly enough. Members of this family are described by

$$S = S_0\left[\cosh\left(\frac{\chi}{h}\right) + T\sinh\left(\frac{\chi}{h}\right)\right]^2 \qquad (18\text{-}70)$$

S_0 is the cross-sectional area at the throat where χ is 0. A scale factor denoted as h is indicative of the rapidity of flare with small values of h corresponding to rapid expansion. T is a shape factor determining the general properties of the horn near the throat. When $T = 0$, Eq. 18-70 generates a catenoidal horn. When $T = 1$, an exponential horn results. When $T = h/\chi_0$ and is allowed to approach ∞ by letting h become very large, a conical horn results. Each of these horns merits individual attention. The descriptions given in each instance assume the horn is long enough so that mouth reflections can be ignored.

Conical Horn

No approximations are involved in the conical horn provided that it is excited at its throat by a spherical wavefront. If mouth reflections are ignored the equations are

$$S = S_0\left[1+\frac{r}{r_0}\right]^2$$

$$p = \frac{p_0}{1+\frac{r}{r_0}} \times e^{j\left(\omega t - \frac{\omega}{c}r\right)} \qquad (18\text{-}71)$$

$$Z_0 = \rho_0 c\left[\frac{1}{1+\left[\frac{c}{\omega r_0}\right]^2} + j\frac{\frac{c}{\omega r_0}}{1+\left[\frac{c}{\omega r_0}\right]^2}\right]$$

In these equations, r is measured from the throat and r_0 is the distance from the virtual apex of the cone to the throat of the horn. It should be observed from Eq. 18-71 that the acoustic pressure varies inversely with distance from the virtual apex of the horn and the phase velocity in the horn is just c. Horns are often compared in terms of their transmission coefficient. The transmission coefficient, τ, is defined to be the ratio of the real part of the throat specific acoustic impedance to the specific acoustic impedance of air for plane waves, this latter value being just $\rho_0 c$. The transmission coefficient behavior for a conical horn is displayed in Fig. 18-25.

Conical horns transmit at all frequencies and thus do not possess a characteristic cut off frequency. Unfortunately, however, the transmission coefficient is small at low frequencies. The pressure attenuation in the horn is the same as that for a spherical wave in free space being a drop of 6 dB for each doubling of distance traveled in the horn.

Cylindrical Horn

A cylindrical horn is constructed by taking a sector from a cylinder of large radius and truncating it at some smaller radius r_0 from the axis of the cylinder so as to provide a throat opening. The top and bottom plates of the horn are flat and separated by a constant distance. The nominal coverage angle of the horn is the angular opening between the vertical plates. The difficulty with this type of horn is the necessity of providing uniform cylindrical wave excitation over the entire throat area. If mouth reflections are ignored the equations are

$$S = S_0\left[1+\frac{r}{r_0}\right]$$

$$p = \frac{p_0}{\left[1+\frac{r}{r_0}\right]^{\frac{1}{2}}} \times e^{j\left(\omega t - \frac{\omega}{c}r\right)} \qquad (18\text{-}72)$$

$$Z_0 = \rho_0 c\left[\frac{1}{1+\left(\frac{c}{\omega r_0}\right)^2} + j\frac{\frac{c}{\omega r_0}}{1+\left(\frac{c}{\omega r_0}\right)^2}\right]$$

The transmission characteristics of a cylindrical horn are identical to those of a conical horn and thus such a horn does not exhibit a cut off frequency. The pressure attenuation within the cylindrical horn, as is true of cylindrical waves in general, is only 3 dB per doubling of distance traveled within the horn.

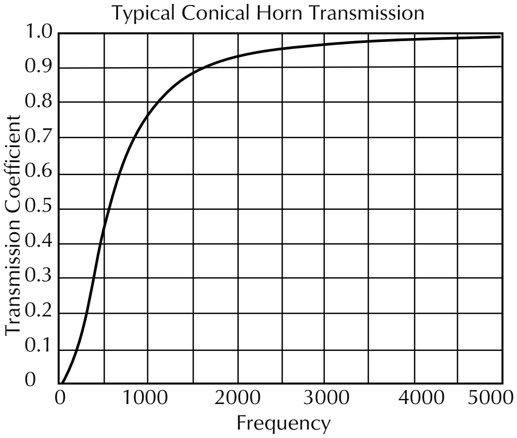


Figure 18-25. Transmission coefficient for a conical horn.

Plane Wave Tubes

Plane wave tubes are used as acoustic delay lines and as aids in testing compression drivers. A plane wave tube is a straight tube of constant diameter d. Such a structure will support only plane waves provided that the frequency of excitation is less than $0.586(c/d)$. If the exciting frequency exceeds this limit, non-planar dispersive modes can be excited in the tube. This limit corresponds to about 8000 Hz for a one inch diameter tube. A semi-infinite plane wave tube would be one which extends along the x-axis say from $x = 0$ to $x = \infty$. Such a tube, when properly excited at $x = 0$, would only support a wave traveling in the direction of increasing x and thus there would be no reflections. Under these conditions, the specific acoustic impedance at the origin of the tube would be just $\rho_0 c$ and the mechanical impedance presented to a driver would be $\rho_0 cS$ where S is the planar cross-sectional area of the tube. Thus, in order to make a tube of finite length appear to be infinite, it must be terminated by a material which has a specific acoustic impedance matching that of air. There are several materials that closely match this requirement including certain grades of fiberglass and acoustic foam. Gradually tapered wedges of such materials usually terminate those plane wave tubes that are intended to be reflection free. The acoustic pressure attenuation in such plane wave tubes depends on the viscous effects of air and is negligible at the frequencies for which such tubes are normally employed. The transmission coefficient of such tubes is thus unity at all frequencies. The equations for such a tube are

$$\begin{aligned} S &= \text{a constant} \\ p &= p_0 e^{j\left(\omega t - \frac{\omega}{c}x\right)} \\ z_0 &= \rho_0 c \end{aligned} \qquad (18\text{-}73)$$

When a plane wave tube of finite length is terminated in a manner which is not reflection free, the mechanical impedance presented by the tube at $x = 0$ depends on both the mechanical impedance at the termination, Z_L, and the length of the tube, L. In this instance the mechanical impedance at $x = 0$ is denoted as Z_0 given by

$$Z_0 = \rho_0 cS\left[\frac{\left(1+\frac{Z_L}{\rho_0 cS}\right)(1+j\tan[kL])}{\left(1+\frac{Z_L}{\rho_0 cS}\right)(1+j\tan[kL])}\right. \qquad (18\text{-}74)$$

$$\left.\frac{-\left(1-\frac{Z_L}{\rho_0 cS}\right)(1-j\tan[kL])}{+\left(1-\frac{Z_L}{\rho_0 cS}\right)(1-j\tan[kL])}\right]$$

Two special applications of Eq. 18-74 occur quite often in practice. Often the tube is terminated by a rigid wall in which instance the particle velocity at the barrier is zero and the mechanical impedance at $x = L$ is then infinite. When Z_L is ∞, Z_0 becomes

$$Z_0 = -j\rho_0 cS[\cot(kL)] \qquad (18\text{-}75)$$

In all instances $k = \omega/c = 2\pi/\lambda$. When the tube is left open at $x = L$, Z_L becomes the complex radiation impedance of a piston of diameter equal to that of the tube with the piston radiating into all of space. From Eq. 18-74, if $L = n\lambda/2$ with n equal to any integer, the tangent terms are identically zero and the mechanical impedance presented by the tube is equal to the termination impedance. An open-ended tube satisfying the half wavelength condition, then, presents a mechanical load to its driving generator equal to the radiation resistance and reactance of a direct radiator whose diameter is equal to that of the tube. If there were some way to cancel the radiation reactance, all of the energy supplied by the generator would be radiated acoustically at the end of the tube. This is indeed possible.

A close examination of Eq. 18-74 reveals that if the tube length is only slightly less than that specified by the half wavelength criterion, then the tube itself will supply the negative reactance necessary to cancel the positive radiation reactance at its end. Under this circumstance the driving generator will see only the radiation resistance presented by the open end. This would constitute the resonant condition for an open-ended tube. Finally, notice from Eq. 18-75 that if $L = n\lambda/4$ where n is an odd integer then the cotangent is zero and the mechanical impedance of the tube becomes zero and the tube can be driven freely at those frequencies for which this is true. This constitutes the resonance condition for a closed end tube.

Catenoidal Horn

The catenoidal horn has a behavior that is quite interesting. In the absence of reflections the equations are

$$\begin{aligned} S &= S_0\cosh^2\left(\frac{x}{h}\right) \\ p &= \frac{p_0}{\cosh\left(\frac{x}{h}\right)} e^{j\left(\omega t - \frac{\omega}{c}\sqrt{1-\left(\frac{c}{\omega h}\right)^2}\,x\right)} \\ Z_0 &= \frac{\rho_0 c}{\sqrt{1-\left(\frac{c}{\omega h}\right)^2}} \end{aligned} \qquad (18\text{-}76)$$

Firstly, this horn exhibits a cut off frequency $f_c = c/2\pi h$ below which true wave motion in the horn ceases to exist. Secondly, unless the operating frequency is well above the cutoff frequency, the phase velocity in the horn is frequency dependent and thus suffers envelope distortion for complex waveforms. A favorable point, however, is that above cut off, the specific acoustic impedance, though frequency dependent, is purely real. Additionally, the slope of the horn walls at the horn throat is zero and this forms a good match to a plane wave tube when the horn is operated well above its cut off frequency. This horn should always be operated well above cut off as the specific acoustic impedance at the throat is infinite at the cut off frequency.

Exponential Horn

The exponential horn is a time-honored device. It performs at its best when the cut off frequency is chosen to be well below the lowest frequency at which the horn is intended to be employed. Neglecting reflections, the equations are

$$S = S_0 e^{\frac{2x}{h}}$$

$$p = p_0 e^{-\frac{x}{h}} e^{j\left(\omega t - \sqrt{\left(\frac{\omega}{c}\right)^2 - \left(\frac{1}{h}\right)^2}(x)\right)} \qquad (18\text{-}77)$$

$$z_o = \rho_0 c\left[\sqrt{1-\left(\frac{c}{h\omega}\right)^2} + j\frac{c}{h\omega}\right]$$

An exponential horn also exhibits a cut off frequency of $f_c = c/2\pi h$. Below the cut off frequency, the throat specific acoustic impedance is purely reactive while above cut off it is part resistive and part reactive. In the transmission band above cut off, the phase velocity is frequency dependent becoming less so as the frequency increases above cut off. It is thus desirable to operate the horn well above cut off in order to reduce group delay distortion and increase transmission. Additionally, the horn comes closer to satisfying the Morse criterion if the cut off frequency is chosen to be very low thus making h large and producing a slowly expanding cross section. Unfortunately, with high throat pressures the non-linear distortion produced in the air is exacerbated by a slowly flaring horn. As is true in many design areas involving loudspeakers, one is often faced with diametrically opposed requirements and compromises must be made. The transmission coefficient for an exponential horn appears in Fig. 18-26.

If one were to apply the Morse criterion to most commercial horns for determining if a given horn structure satisfies the conditions necessary for Webster's equations to be valid, the results would be a cause of great dismay. Recall the Morse criterion requires the rate of growth of the square root of the cross-sectional area with distance must be much less than one. With the possible exception of a single cell in a multi-cellular horn having a very low cut off, none of the common exponential horns come even close to satisfying the slowness of growth criterion. This does not mean that the horns do not work. It simply means that the solutions of Webster's equations do not adequately describe how they work. The actual wave motion in the horn is much more complex than is assumed and involves a variety of modes as well as internal reflections.

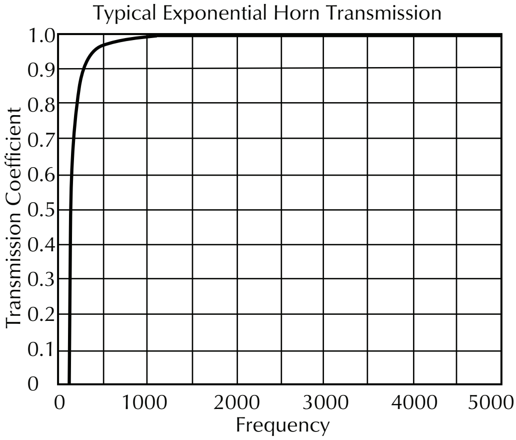


Figure 18-26. Transmission coefficient of a 125 Hz exponential horn.

With the exception of the cylindrical horn, all of the early horns were figures of revolution about the principal horn axis and thus had circular cross sections. This is not a necessary constraint as the required area growth with distance within the horn can be obtained with other cross sectional shapes. Most modern horns have rectangular cross sections so as to produce coverage patterns that are wider horizontally than vertically. This shape more nearly conforms to the general audience-seating pattern. Most if not all horn drivers have circular exit apertures which means that horns that are ultimately rectangular must be circular at the throat and hence must include a smooth transitional section. If the horn walls near the mouth are to approximately determine the coverage angles of the horn, however, very narrow mouths are to be avoided as mouth diffraction will greatly broaden the coverage angle in what was to be the narrow direction.

Horns and horn drivers cannot be treated as separate entities without the payment of penalties in the resulting overall performance. For example, a cylindrical tube needs to be driven by a piston as does a catenoidal horn whereas a conical horn requires spherical excitation. If the chosen driver does not supply the excitation that matches the horn throat, a reflection will occur at the juncture of the horn throat with the driver. Additionally, in many compression drivers, area expansion begins within the phase plug of the driver. In effect then, the horn actually begins in the phase plug. The point is that the driver and the horn with which it is to be used should be designed as a system rather than as separate entities.

All of the horns so far, other than conical and cylindrical, exhibit behavior that departs from that predicted by the simple theory particularly at higher frequencies. As long as the wavelength is large compared with the horn diameter of the horn at any point, the conditions required by the simple theory are approximately satisfied and operation is close to prediction. When the wavelength becomes comparable to or less than the horn diameter at any point the wave will no longer fill the entire horn cross section and the beam width is no longer determined by the horn geometry. This is a diffraction-related phenomenon in which the acoustic energy is concentrated in an increasingly narrower lobe about the horn axis as the frequency increases. The coverage angle of such a device narrows as the frequency is increased. This behavior has been circumvented in recent years by a different horn design technique that leads to a horn having a more nearly frequency independent coverage pattern.

Constant Directivity Horns

A constant directivity horn is a hybrid design incorporating aspects from several other horn types. A typical design begins with a throat section having an exponential taper with a cut off frequency well below the lowest operating frequency for which the horn is to be employed. This assures an adequate acoustic load for the horn driver. This throat section furnishes additionally a smooth transition from a circular cross section to a skinny rectangular cross section such that the exit from the throat section is in the form of a slot.

The wave fronts in the interior of the throat section have very little curvature. When these waves encounter the slot two different events occur. A portion of the wave is reflected back towards the driver as the exit from the throat section represents a discontinuity in the wave medium. The portion of the wave which emerges from the slot undergoes diffraction at the edges of the slot and emerges as a diverging wave such that the wavefront has a small radius of curvature in a plane perpendicular to the length of the slot and a larger radius of curvature in a plane parallel to the length of the slot.

The diverging wave is allowed to expand into a bounded region whose cross-sectional area expands conically except that the horizontal angle is greater than the vertical angle so as to match the differing radii of curvature formed at the slot. Wave propagation in the expanding conical section is well behaved as the horn walls provide a reasonable match for the diverging wavefronts. The wavefronts, of course, are astigmatic by virtue of the differing radii of curvature in the horizontal and vertical planes. With a sufficiently large mouth, the horn coverage angles will be independent of frequency up to a limit where the wavelength becomes comparable to the width of the slot. Beyond this limit the horn will begin to beam. In spite of the constant directivity property, horns of this type have two distinct and serious drawbacks. The internal reflection that occurs at the diffraction slot gives rise to standing wave production with associated resonances occurring in the throat section. The astigmatism associated with the emerging wavefronts from the horn implies separated sources of sound rather than a single point source, as would be the case for a purely spherical wavefront. This makes it difficult to combine two or more horns so as to have enlarged coverage. Such horns require one placement for a horizontal combination and a completely different placement for a vertical combination.

Horn theory is still a work in progress. In the mid 1970s Benade described the wave motion in axi-symmetric horns of conventional shapes in terms of spherical waves possessing both radial and axial modes. A large body of experimental measurements illustrating the usefulness of this approach accompanied his theoretical work. Recently Geddes has proposed new horn shapes based on oblate spherical coordinates.

18.10 Practical Considerations Involving Horns

Horns are seldom employed as single stand alone devices except in simple paging systems. More often they are employed as mid or high frequency devices in conjunction with vented enclosure type low frequency direct radiator devices in forming a full range loudspeaker. Additionally, they are often employed in large two-way systems where they are arrayed either to increase coverage or pressure level or both. In this instance, low frequency support is provided by a separate array of low frequency devices.

In the early 1980s a full range loudspeaker design was introduced that at low frequencies performed as a vented enclosure, in the mid frequencies was loaded by a large mouth horn, and for the high frequencies featured a co-axially mounted high frequency horn located in the interior of the mid frequency horn. The coverage angles, acoustic origins, and acoustic centers of the horns were similar enough that the combination acted as a single device and thus could be readily arrayed with similar devices. Since that time many manufacturers have developed similar full range loudspeakers. The current trend is to build arrays based on such full range devices. In developing and subsequently arraying such devices knowledge of common horn properties must be available. These properties are sensitivity, directional *Q*, acoustic origin, acoustic center, loudspeaker efficiency, and polar response. Sensitivity, *Q,* and efficiency have been defined previously. The other terms require definition and further discussion.

If one measures the impulse response of a loudspeaker at some remote point in space, it is found that a finite time elapses between the application of the loudspeaker excitation and the reception of the response. This time interval is called the transit time. The product of the transit time with the value of the speed of sound that prevails under the conditions of measurement defines a distance.

When this distance is projected from the observing microphone towards the loudspeaker along the loudspeaker axis, the terminus of this projection defines a point in space from which the signal appears to originate. This point is called the acoustic origin. On the other hand, the acoustic center is the point in space from which the spherical wavefronts appear to diverge as observed in the far field. Fig. 18-27 illustrates the location of these points for a variety of horns.

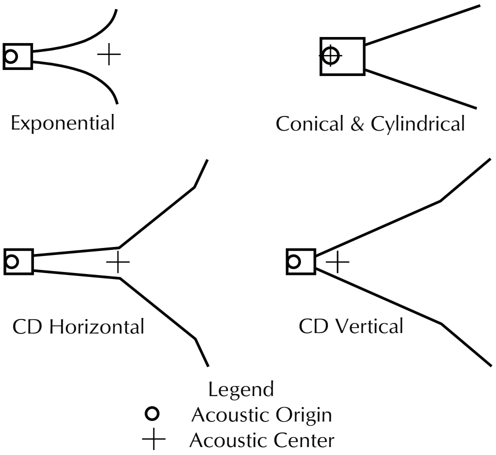


Figure 18-27. Typical locations of acoustic origins and acoustic centers.

In viewing Fig. 18-27, it should be observed that only the cylindrical and conical horns have co-incident locations for both the acoustic origin and acoustic center.

The polar response of a loudspeaker, whether the loudspeaker is a full range system or an individual device, is a quantitative way of describing the loudspeaker's directional characteristics. It is in effect a radiation pattern. One is accustomed to viewing manufacturer's condensed specification sheets wherein one often encounters what are called horizontal and vertical polars. Such plots are useful but by no means do they constitute a complete description of a device's directional characteristics.

In order to obtain a complete description it is necessary to make an extensive series of measurements in the far field of the device. In making these measurements, it is necessary to excite the device with various bands of pink noise as the polar response is frequency dependent. Additionally, pressure level observations must be made versus the angular orientation of the device. The angles involved are the axial angle that measures rotation about the device's principal axis and the azimuthal angle that measures rotation about the device's acoustic center. These rotations are illustrated in Fig. 18-28.

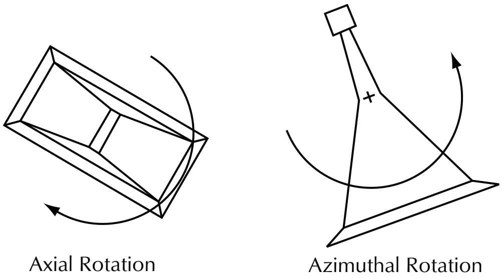


Figure 18-28. Rotational axes for polar measurements.

There is not a unique procedure that must be followed in making polar measurements. What will be described is a possible procedure that will accumulate the required data. The measurement session starts with both the axial and azimuthal angles set at zero with the measuring microphone located on the principal axis of the device in the far field. One then excites the device sequentially with the pertinent bands of pink noise. These noise bands should be no wider than one octave.

The pressure level on axis for each noise band is separately recorded. The device is then rotated about the azimuthal axis through a fixed angular increment such as five degrees. Again the sequence of noise bands is applied with the corresponding pressure levels being recorded. The azimuthal angle is again incremented and the pressure levels are recorded appropriately.

One proceeds in this fashion until the device is swept through one complete revolution about the acoustic center. What results are the horizontal polars. Now starting again with the azimuthal angle back at zero degrees, one now increments the axial angle by a fixed increment such as five degrees and then repeats the previous set of observations wherein the azimuthal angle is again swept through a complete revolution. At the conclusion of this second sweep, one again increments the axial angle by an additional five degrees. Again one makes observations as the azimuthal angle is swept through one complete revolution. This process continues until the axial increments total 90° for a device that has mirror symmetry in a plane containing the principal axis.

For other devices, the process must continue until the axial increments sum to 180°. The sweep where the axial increments sum to 90° generates the vertical polars. The pressure levels thus obtained usually are normalized to the values obtained on the principal axis with zero azimuth and zero axial angles. For each band of pink noise, one arrives at an m by n array of pressure values where m is the number of azimuthal angle positions and n is the number of axial angle positions. These values constitute the polar response at least to the precision that is set by the size of the angular increments employed. Polar response data in this or similar form is required for use in computer generated loudspeaker coverage plots.

The foregoing procedure suffers in two respects. Firstly, only magnitude response data is accumulated and secondly, the measurement microphone must truly be in the far field for accurate results. Most facilities with the controlled environments necessary for unperturbed measurements are not large enough to allow far field measurements on large devices at high frequencies. Measurements in such facilities are usually made at distances in the transition region between the near and far fields and have inaccuracies as a result.

Gunness and Mihelich, however, have introduced techniques for accurately predicting far field directional response based upon near field measurements made on a geometrical surface surrounding the loudspeaker under test. The adoption of such procedures will lead to more accurate descriptions of loudspeaker directional characteristics of both magnitude and phase in the future.

On a final note, there is a useful relationship between loudspeaker sensitivity, loudspeaker axial Q, and loudspeaker electrical efficiency coefficient. This relationship appears as Eq. 18-78.

$$\eta = \frac{1}{Q}10^{\left(\frac{L_p}{10} - 10.9\right)} \tag{18-78}$$

where,

η is the efficiency coefficient,

L_p is the 1 watt @ 1 m sensitivity,

Q is the axial directivity factor.

When employing Eq. 18-78 it is important to remember that each of the variables is frequency dependent. It is necessary that the values employed must be those for a given frequency or for a given band of frequencies.

18.11 Horn Compression Drivers

A typical structure for a compression driver appears in Fig. 18-29.

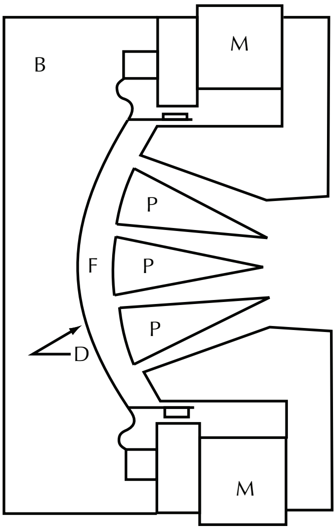


Figure 18-29. Typical compression driver displayed in cross section. M is the magnet, B is the back cavity, D is the diaphragm, F is the front cavity, and P is the phase plug.

The view presented in Fig. 18-29 is a sectional view in the median plane of the driver. The magnet, which is actually in the form of a ring, is sandwiched between the pole pieces. The pole pieces define the air gap in which the voice coil resides. A portion of the front pole piece also forms part of the phase plug. The depicted phase plug has concentric annular slits as openings to expanding air channels. These channels provide air paths of equal length as measured from the rear face of the phase plug. The rear face of the phase plug and the front face of the diaphragm define the front air cavity. The rear face of the diaphragm and the rear cover of the driver define the back air cavity.

Area expansion actually begins in the phase plug and continues up to the driver's front face. An attached horn is ideally just an extension of this expansion. The horn behavior, then, actually begins in the phase plug. The compression ratio of such a driver is the ratio of the front surface area of the diaphragm to the entrance area of the air channels in the phase plug. The area of the diaphragm is taken to be S_D while that of the phase plug entrance area is S_T. The compression ratio is thus S_D/S_T. This ratio typically has a value of 10. The electrical and mechanical behavior of a compression driver is similar to that of a back enclosed direct radiator as far as the back cavity of the driver is concerned. The loading on the front face of the diaphragm, however, is markedly different from that of a direct radiator. On its front face, the diaphragm must compress or expand the air in a cavity of small volume with relief only being supplied through air flow in the restricted entrance area of the channels in the phase plug. The appropriate model accounting for this behavior is presented in Fig. 18-30.

The model of Fig. 18-30 is drawn for the circumstance where the driver is loaded by a plane wave tube or an ideal horn operating well above its cut off frequency. This load is coupled into the driver through an ideal transformer whose turns ratio is just the compression ratio of the driver. The mechanical impedance acting directly on the back side of the diaphragm is Z_D.

$$Z_D = R_m + R_r + j\left[\omega M_D + X_r - \frac{\left[K_D + \dfrac{\rho_0 c^2 S_D^2}{V_B}\right]}{\omega}\right] \tag{18-79}$$

where,

R_m is the mechanical resistance of the diaphragm,

R_r is the radiation resistance on the back side of the diaphragm,

M_D is the diaphragm effective mass,

X_r is the radiation reactance on the back side of the diaphragm,

K_D is the diaphragm suspension stiffness,

V_B is the back cavity volume.

The mechanical impedance acting directly on the front side of the diaphragm is Z_F.

$$Z_F = \frac{\dfrac{\rho_0 c^2 S_D^2}{j\omega V_F}\rho_0 c S_T}{\dfrac{\rho_0 c^2 S_T^2}{j\omega V_F} + \rho_0 c S_T} \tag{18-80}$$

where,

V_F is the front cavity volume.

An analysis of the "circuit diagram" provided by the model leads to expressions for the velocity of the diaphragm, u_D, and the air particle velocity at the entrance to the phase plug, u_T. The velocity of the diaphragm is

$$u_D = \frac{\dfrac{E_S Bl}{R_e + j\omega L_e}}{\dfrac{(Bl)^2}{R_e + j\omega L_e} + Z_D + Z_F} \tag{18-81}$$

The air particle velocity at the entrance to the phase plug is the "current" in the ideal transformer secondary and is

$$u_T = \frac{\left(\dfrac{S_D}{S_T}\right)\dfrac{\rho_0 c^2 S_T^2}{j\omega V_F}}{\dfrac{\rho_0 c^2 S_T^2}{j\omega V_F} + \rho_0 c S_T} u_D \tag{18-82}$$

The detailed solution of these last two equations is the province only of the driver design engineer as numerical values of the driver parameters are not generally available. Nevertheless a few general conclusions may be formed. If one considers the frequency range in the vicinity of the middle of the pass band of the driver, the total impedance in the denominator of Eq. 18-81 can be made almost purely resistive.

This occurs as follows: In this frequency range the reactance of the front cavity acts as an open circuit so that Z_F is then purely resistive with a value of $(S_D/S_T)^2 \rho_0 c S_T$. If now Z_D is tuned to resonance in this frequency range, this resonance will necessarily be of low Q because of the dominating large resistive terms. As a consequence, there will exist a broad frequency range where u_D will be constant and the power response of the driver will be flat. Above this frequency range, the positive mass reactance of the diaphragm will begin to dominate and the power response will fall at the rate of 6 dB/octave. This will continue until the frequency becomes high enough that one can no longer ignore the front cavity reactance. In this range, one has the front cavity negative reactance acting in parallel with the resistive load on the driver with this combination in series with the positive mass reactance of the diaphragm. This combination now constitutes a second order low pass filter forcing the power response to now fall at 12 dB/octave.

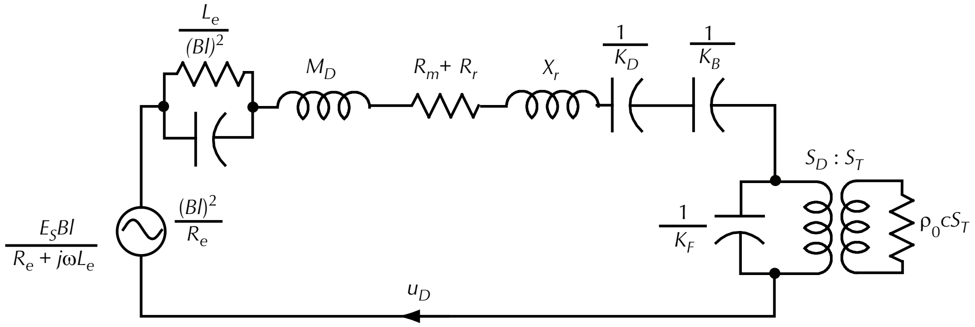


Figure 18-30. Mechanical impedance model of a compression driver.

Eventually, however, the voice coil inductance will become important, forcing the power response to ultimately fall at 18 dB/octave. This behavior is accurately accounted for in the model, but the model is ignorant of breakup modes that occur in the diaphragm. The artifacts of these modes usually are superimposed in the region where the model predicts a fall of 12 dB/octave.

In this region there will be response peaks and valleys as a result of resonances occurring in the diaphragm material itself. There is one alternative construction that can force a modification of the behavior described above. If the front cavity volume is made very small indeed, it is possible to force a series resonance between the negative reactance of this cavity and the positive mass reactance of the diaphragm in the region where the response would have previously been falling at 6 dB/octave.

This resonance will also be of low Q and will extend the range of driver flat power response to higher frequencies than before. This extended response will come with a large penalty. In making the front cavity volume very small, one has seriously restricted the maximum allowable displacement of the diaphragm and consequently the maximum driver power output. This is another instance where a trade off must be made in the design process. Typical alternative behaviors predicted by the model are the two curves appearing in Fig. 18-31.

The measured performance of driver-horn combinations can and does depart markedly from that depicted in Fig. 18-31. Firstly there are artifacts of diaphragm breakup previously mentioned and secondly, there are the effects of resonances from both mouth and internal reflections occurring in the attached horn. The overall performance is considerably less smooth than that implied by Fig. 18-31. As an illustration of this fact, Fig. 18-32 is the measured response of a compression driver mounted on a constant directivity horn.

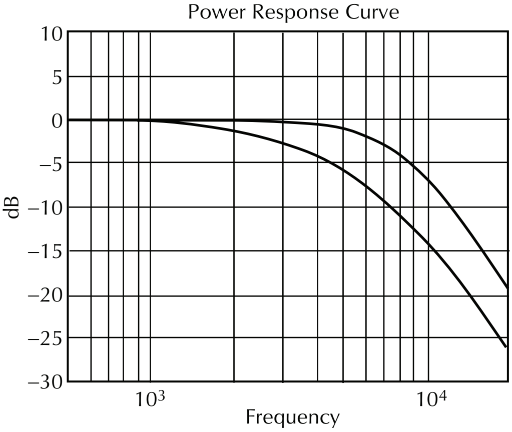


Figure 18-31. Theoretical high frequency performance of a compression driver.

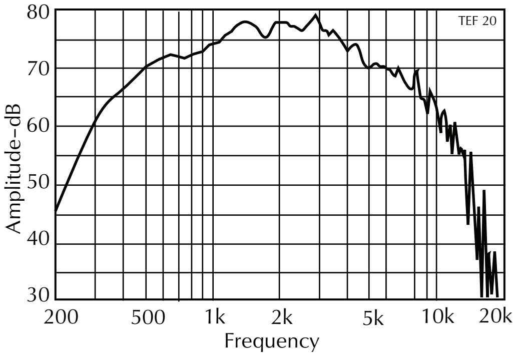


Figure 18-32. Compression driver on a constant directivity horn.

18.12 Crossover Networks

Origin

Crossover networks or frequency dividing networks have their origin with the development of two-way loudspeakers in the motion picture industry during the decade of the 1930s. The early networks were passive systems interposed between the single power amplifier typical of the time and the high and low frequency stage loudspeakers positioned behind the perforated projection screen. These networks were either constant-k or combinations of constant-k with m-derived sections based on the filter technology that had been developed by Bell Telephone Laboratories.

In the 1940s and 1950s constant resistance crossover networks made their appearance based upon the newer filter technology which involved Laplace transforms and pole-zero analysis. In the 1950s and 1960s active crossover networks made their initial appearance with ever increasing importance being displayed in the 1970s attendant with the growth of integrated circuits. Increasing attention was also being given at this time to phase behavior, driver physical positioning, and driver phase response. The 1980s presented not only a well-developed body of theory for network design but also new acoustical measuring instruments for assessing overall system performance in the acoustical domain. This work continues to the present day.

Electric Filters

Electric filters are circuit assemblies which modify a signal's properties as a function of frequency. This modification may involve amplitude, phase, or both. One category of filter freely transmits signals over certain ranges of frequency while attenuating or failing to transmit in other ranges. This attenuation is accompanied by a frequency dependent phase shift. Minimum phase shift filters fall in this category. There is another category of filter which transmits freely at all frequencies without affecting signal amplitude but rather introduces a frequency dependent phase shift. All pass filters fall into this category and they find application in modern crossover networks both active and passive.

Transfer Functions

In order to understand crossover networks and, indeed, even to be able to read current literature on the subject it is necessary to have some knowledge of transfer functions. The relationship which exists in the steady state between the output signal and the input signal of a two port device such as a filter is called the transfer function. The transfer function has a magnitude and an angle with each, in general, being frequency dependent. The magnitude of a transfer function is an expression of the ratio of the output signal amplitude to the input signal amplitude. The angle of a transfer function is the phase difference between the output signal and the input signal. This is best illustrated by means of a simple example. Fig. 18-33 depicts a simple low pass filter consisting of an inductance and a resistance.

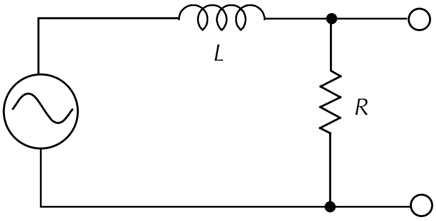


Figure 18-33. Simple low pass filter.

Through the employment of the techniques of ac circuit analysis, one may write the relationship between the output voltage and the input voltage of the circuit of Fig. 18-33 as

$$H = \frac{\frac{R}{L}}{j\omega + \frac{R}{L}} \tag{18-83}$$

Alternatively, this may be written in the complex exponential form

$$H = \frac{\frac{R}{L}}{\sqrt{\omega^2 + \left(\frac{R}{L}\right)^2}} e^{j \operatorname{atan}\left(\frac{-\omega}{\frac{R}{L}}\right)} \tag{18-84}$$

and finally as

$$H = \frac{\omega_0}{S + \omega_0} \tag{18-85}$$

where,

H is the complex transfer function,

S is the Laplace transform variable whose steady state value is $j\omega$,

$\omega_0 = R/L$.

The form presented in Eq. 18-85 is the statement of the transfer function of the example low pass filter in the language of the Laplace transform. In the general case which includes the transient state, S can have both real and imaginary parts. In this case, S can be written as $S = \sigma + j\omega$. S is called the complex frequency variable. If S is allowed to assume any possible value whether it be real, imaginary, or complex such that all points in a two dimensional complex plane are accessible, there would be only one value of S for which the denominator of Eq. 18-85 would become zero.

This value of S is $S = -\omega_0$. When S takes on this value, the denominator becomes zero and the magnitude of the transfer function becomes infinite. It is said then that this transfer function has a single "pole" located at $S = -\omega_0$. The pole order of a transfer function is determined by the highest power of S appearing in the denominator.

A two pole filter would have an S^2, a three pole would have an S^3, etc., appearing in the denominator of the transfer function. In the steady state as opposed to the transient state S is restricted to the values $S = j\omega$ and the only accessible points lie on the positive imaginary axis. In the steady state, even though the value of S never coincides with the location of the example pole, the pole location nevertheless influences the operation of the filter.

Changing the pole location changes the value of ω_0 and hence changes the value of the transfer function at all frequencies other than zero. A further study of the Laplace transform and the inverse Laplace transform indicates that the transfer function is also a description of the filter's impulse response in the complex frequency plane. Additionally, the inverse Laplace transform of the transfer function is the description of the filter's response to an impulse described in the time domain, i.e., it is the filter's transient response to an impulsive excitation expressed as a function of time.

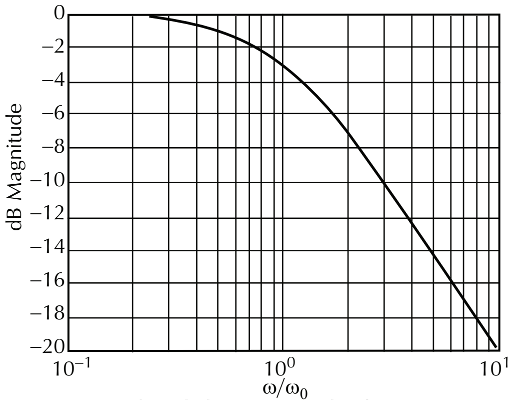
An important consequence of this is that in order for a filter to exhibit a transient response that decays with increasing time, all of the poles of the filter's transfer function must have negative real parts. The filter under discussion satisfies this criterion as its single pole is located at $-\omega_0$, and hence its transient response decays with time thus allowing the filter to exhibit a stable steady state response.

The information contained in a filter's transfer function may be depicted in a variety of ways, the two most popular of which are the Bode plot and the Nyquist diagram. In exploring this it should be noted that Eq. 18-84 is of the form

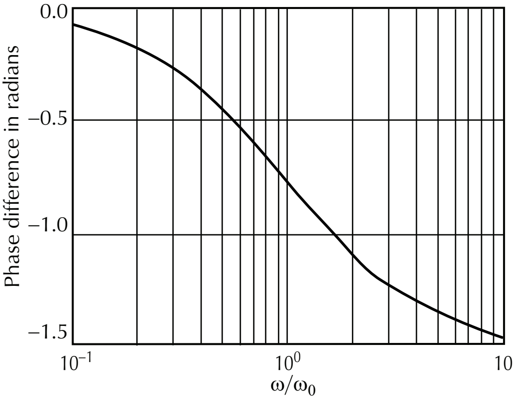
$$H = |H|e^{j\varphi} \tag{18-86}$$

The Bode plot displays Eq. 18-86 in the form of a graph of $20\,\text{dB}\log|H|$ versus $\log(\omega)$ and a graph of φ versus $\log(\omega)$. This behavior is displayed in Figs. 18-34A and 18-34B.

In Fig. 18-34A, the pass band ranges from zero frequency to the point where $\omega = \omega_0$ where the attenuation has become −3 dB. Note that in the stop band (the region of increasing attenuation beyond the −3 dB point) the slope approaches −20 dB/decade or equivalently −6 dB/octave.



A. Single pole low pass amplitude response.



B. Single pole low pass phase response.

Figure 18-34. Low pass filter response.

Fig. 18-34B suggests that the phase shift is zero at zero frequency, $-\pi/4$ at $\omega = \omega_0$, and approaches $-\pi/2$ at high frequencies.

The behavior of this single pole low pass filter is displayed in a different manner by means of a Nyquist diagram. A Nyquist diagram is a graph of the real part of a frequency dependent complex quantity versus the imaginary part of the same complex quantity. In making this graph, the frequency is treated as a parameter and is allowed to range from zero to infinity. Fig. 18-35 is the Nyquist diagram for the single pole low pass filter.

The same information displayed in the Bode plot is also contained in the Nyquist diagram. The magnitude of the transfer function at any frequency corresponds to the length of a line drawn from the origin to the point on the curve in Fig. 18-35 which corresponds to the frequency value in question. The angle of the transfer function at any frequency is the angle between the magnitude line described above and the real axis.

Corresponding to the first order low pass filter of the previous example is a complementary first order high pass filter. The circuit for this filter appears in Fig. 18-36.

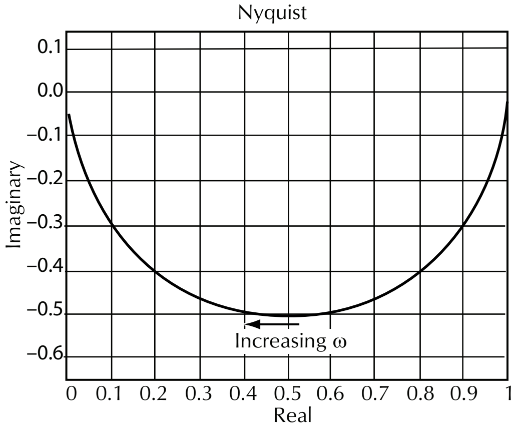


Figure 18-35. Nyquist diagram for a single pole low pass filter.

The high pass filter equations written in the same order as before are

$$H = \frac{R}{R + \frac{1}{j\omega C}} = \frac{j\omega}{j\omega + \frac{1}{RC}} \tag{18-87}$$

$$H = \frac{\omega}{\sqrt{\omega^2 + \left(\frac{1}{RC}\right)^2}} e^{j\left[\frac{\pi}{2} - \operatorname{atan}(\omega RC)\right]} \tag{18-88}$$

$$H = \frac{S}{S + \omega_0} \tag{18-89}$$

where,
$\omega_0 = 1/RC$.

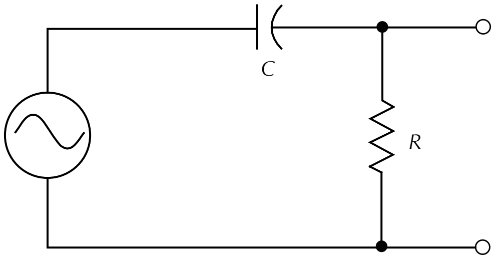
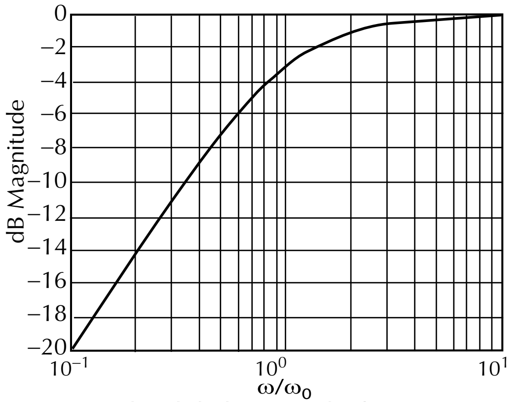


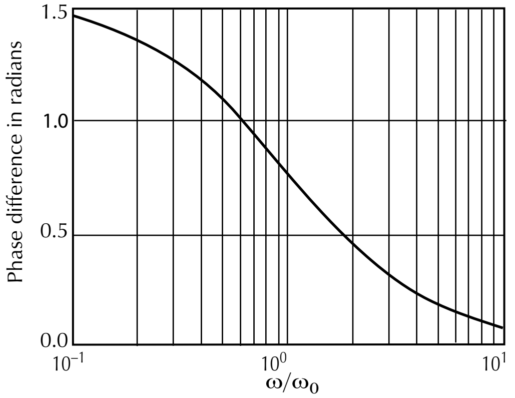
Figure 18-36. Simple high pass filter.

The performance of this filter is depicted in the Bode plots of Figs. 18-37A and 18-37B as well as the Nyquist diagram of Fig. 18-37C.

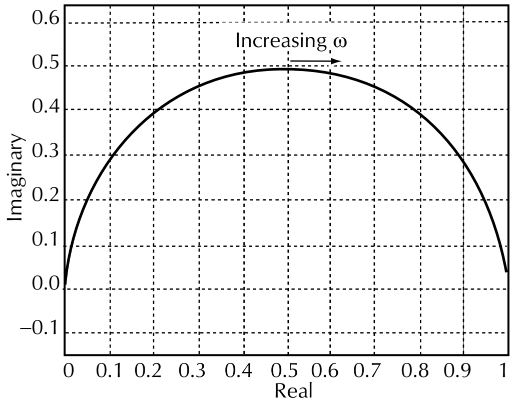
Upon returning to Eq. 18-89, it is apparent that this transfer function has a pole located at $\omega = -\omega_0$.



A. Single pole high pass amplitude response.



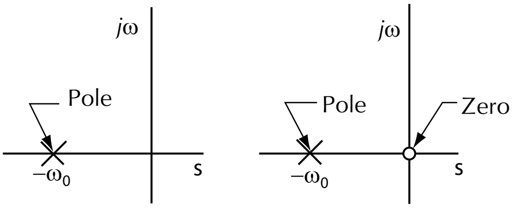
B. Single pole high pass phase response.



C. Nyquist diagram for a single pole high pass filter.

Figure 18-37. High pass filter response.

In addition there is an S appearing in the numerator of the equation. In the complex plane, if there is a value of S which makes the numerator of a transfer function equal to zero, such a value is called a "zero" of the transfer function. A pole-zero diagram is a useful way of displaying this information. Fig. 18-38A is the pole-zero diagram for the first order low pass filter while Fig. 18-38B is that for the corresponding high pass filter.



A. Low pass. B. High pass.

Figure 18-38. Pole-zero diagram.

When one superimposes the amplitude response curves of the low pass and high pass filter as is done in Fig. 18-39, it is found that the two curves cross each other at $\omega = \omega_0$. This behavior is the origin of the term crossover network. At this point each of the terminating resistors is receiving ½ of the total power supplied.

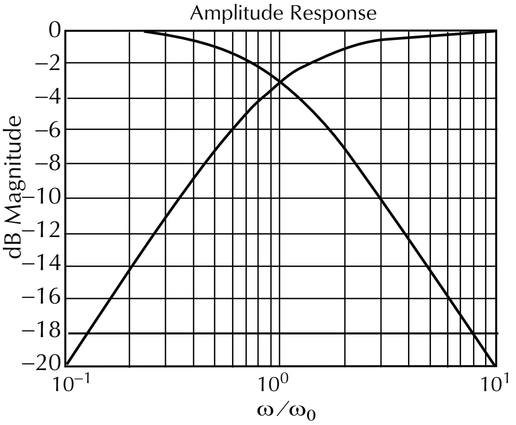
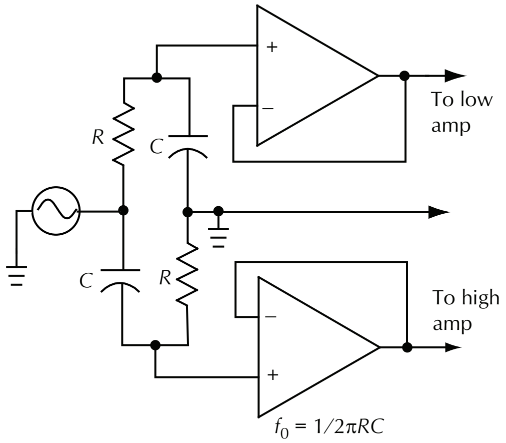


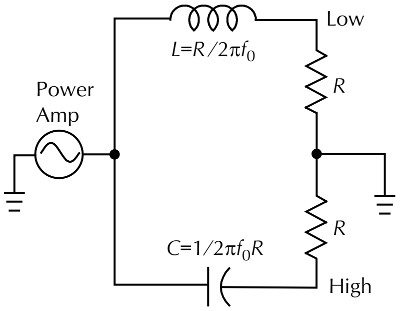
Figure 18-39. Illustration of crossover point for single pole low and high pass filters.

This pair of filters could constitute a simple crossover network for a two-way loudspeaker system. The implementation can be accomplished either actively or passively. The active implementation can be accomplished in either the analog or digital domain. A discussion of the digital implementation appears in Chapter 22 *Signal Processing*. Examples of active analog and passive implementations are shown in Fig. 18-40.

In order for the passive version to work properly, the terminating loads must be constant resistances at least for an octave below and an octave above crossover. These loads are represented by *R* in the circuit diagram. In practice the loads are the impedances of the low and high frequency loudspeakers. Moving coil loudspeakers, as displayed in Fig. 18-20, have impedances which are far from constant, nor are they purely resistive. As is shown in Chapter 8 *Interfacing Electrical and Acoustic Systems*, Section 8.1, Alternating Current Circuits, however, it is possible to parallel the loudspeaker with a Zoebel network which will produce a combination having a sufficiently constant resistance provided that the chosen crossover frequency is well removed from the driver's mechanical resonance frequency. When one calculates the electrical impedance presented to the power amplifier in the passive version of the network in Fig. 18-40 while employing the component values indicated, the impedance is found to be *R* at all frequencies. This network, then, falls into a category termed as being a constant resistance network. The filters in this simple network possess many other interesting and useful properties. These filters have an amplitude response that is maximally flat which means they are Butterworth filters. As each of the filters has a single pole, they are first order Butterworth filters. If one sums the transfer functions of the low pass and the high pass as is done in Eq. 18-90 it is found that the sum is unity. This implies that the combination forms a special all pass filter having a constant amplitude response at all frequencies with zero phase shift at all



A. Active network.



B. Passive network.

Figure 18-40. Active and passive versions of first order crossover.

frequencies. This means that the summed response would exactly replicate the original input signal.

$$\frac{\omega_0}{S+\omega_0} + \frac{S}{S+\omega_0} = 1 \tag{18-90}$$

Additionally, the power response of a filter is given by the square of the absolute magnitude of the filter's transfer function. If one adds the power responses of the filters comprising this simple first order crossover network, as is done in Eq. 18-91, it is found that the combined power response of this network is also unity.

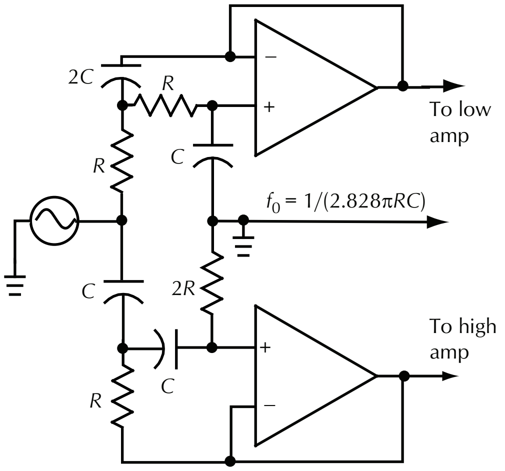
$$\frac{\omega_0^2}{\omega^2+\omega_0^2} + \frac{\omega^2}{\omega^2+\omega_0^2} = 1 \tag{18-91}$$

This unity power summation is a property that is shared, in fact, by all orders of Butterworth two-way crossover networks. Uniform power response generally leads to improved intelligibility in semi-reverberant and highly reverberant fields. This fact is responsible for the popularity of Butterworth crossovers for employment in speech reinforcement systems.

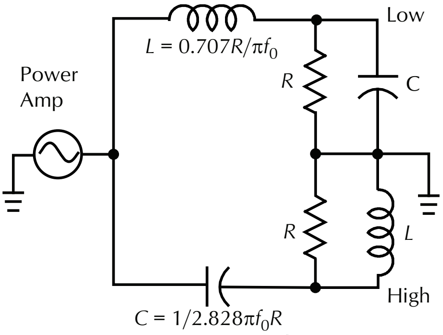
In summary, this first order Butterworth two-way crossover network appears to be ideal. The network is simple in construction, is maximally flat, has unity summed amplitude response with no frequency dependent phase shift, and has an overall unity power response. In fact, it is the only two-way network of any order possessing all of these desirable properties. This network's sole shortcoming is the modest attenuation rate of 6 dB/octave either side of crossover. This small attenuation rate does not adequately protect high frequency drivers against low frequency signals except in low power systems where the network may be employed quite successfully. In high power systems higher order networks are usually required.

Higher Order Networks

The need for greater attenuation slope leads immediately to filters and networks of second or higher order. For many years the "standard" crossover network of the constant resistance variety was the second order Butterworth. This type was chosen because of its increased attenuation rate, maximally flat amplitude response, and unity power summation property. The active and passive realizations of the second order Butterworth crossover are given in Fig. 18-41.



A. Active network.



B. Passive network.

Figure 18-41. Second order Butterworth crossover.

The transfer function describing the low pass section behavior is

$$H_l = \frac{\omega_0^2}{S^2 + \sqrt{2}S\omega_0 + \omega_0^2} \tag{18-92}$$

while that for the high pass section is

$$H_h = \frac{S^2}{S^2 + \sqrt{2}S\omega_0 + \omega_0^2} \tag{18-93}$$

The Bode plots for these functions are given in Figs. 18-42 and 18-43.

The increased attenuation slopes of 12 dB/octave are evident in Fig. 18-42. When one examines the summed frequency response as given in Eq. 18-94, however, it is found that all is not well.

$$H_l + H_h = \frac{S^2 + \omega_0^2}{S^2 + \sqrt{2}S\omega_0 + \omega_0^2} \tag{18-94}$$

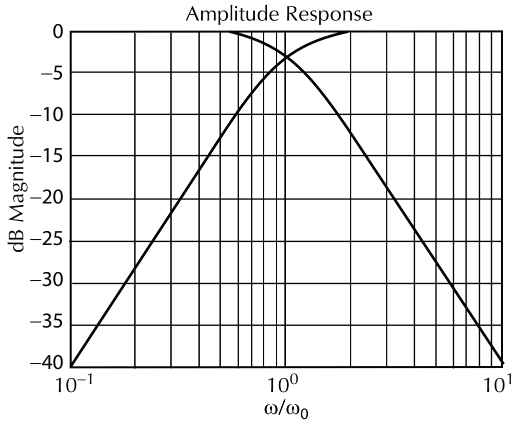


Figure 18-42. Magnitude response of second order Butterworth crossover.

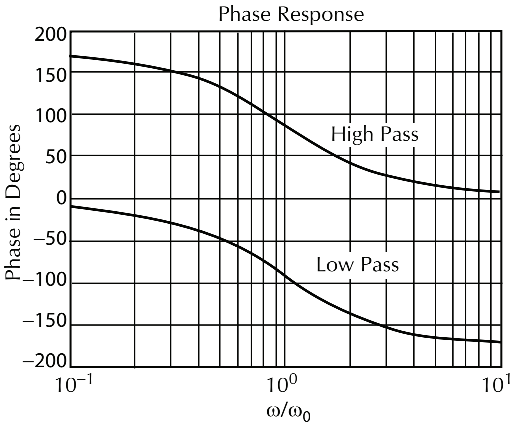


Figure 18-43. Phase response of second order Butterworth crossover.

The summed frequency response is not independent of frequency as it was in the first order case. This summed response has a value of unity at the frequency extremes when S is either very small or very large while at the crossover frequency, where S is $j\omega_0$, the summed response is zero! The reason for this is graphically displayed in Fig. 18-43. The response magnitudes of the low pass and high pass sections are equal at the crossover but the phase of the high pass is +90° or $\pi/2$ while that of the low pass is –90° or $-\pi/2$ and this phase difference leads to a net sum of zero.

This behavior led one audio company to include a third driver in their basically two-way loudspeaker system to “fill the hole in the middle.” Much advertising copy was generated by this approach. Another approach, taken by others, is to reverse the polarity of the driver in the low frequency section of a two-way loudspeaker system. This has the effect of taking the difference between H_h and H_l for the summed frequency response as given in Eq. 18-95.

$$H_h + H_l = \frac{S^2 - \omega_0^{\;2}}{S^2 + \sqrt{2}S\omega_0 + \omega_0^{\;2}} \tag{18-95}$$

Eq. 18-95 has a value of −1 at very low frequencies, a value of 1.414j at crossover, and a value of +1 at very high frequencies. The corresponding values in the language of the Bode plots would be 0 dB with an angle of π at low frequencies, +3 dB with an angle of $\pi/2$ at crossover, and 0 dB with an angle of 0° at high frequencies. This technique greatly improves the summed amplitude response while requiring an overall phase change of π radians between the frequency extremes. Regardless of which of these two possible connections are employed, the summed power responses will still be unity.

An examination of the third order Butterworth network will enable some general statements about subsequent higher order networks. The transfer functions for the third order Butterworth are

$$H_l = \frac{\omega_0^{\;3}}{S^3 + 2S^2\omega_0 + 2S\omega_0^{\;2} + \omega_0^{\;3}} \tag{18-96}$$

and

$$H_h = \frac{S^3}{S^3 + 2S^2\omega_0 + 2S\omega_0^{\;2} + \omega_0^{\;3}} \tag{18-97}$$

The Bode plots for this network appear in Figs. 18-44 and 18-45.

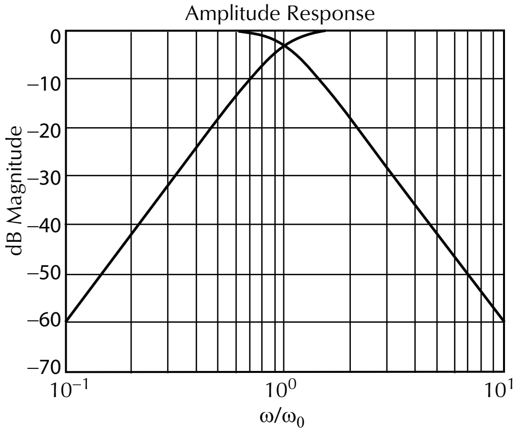


Figure 18-44. Magnitude response of third order Butterworth crossover.

From Fig. 18-44 it should be apparent that the attenuation slopes are now 18 dB/octave or equivalently 60 dB/decade. From Fig. 18-45 it is seen that the phase shift in each filter between frequency

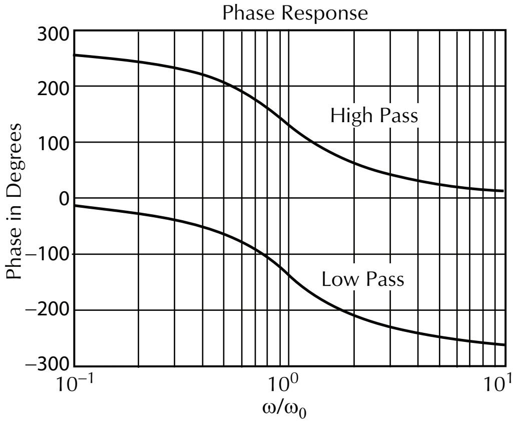


Figure 18-45. Phase response of third order Butterworth crossover.

extremes is 270° with exactly half of the overall having been accomplished at the crossover frequency. In each instance these values are just three times the corresponding values for the first order network. The summed frequency response of the third order Butterworth network appears in Eq. 18-98.

$$H_l + H_h = \frac{S^3 + \omega_0^{\ 3}}{S^3 + 2S^2\omega_0 + 2S\omega_0^{\ 2} + \omega_0^{\ 3}} \qquad (18\text{-}98)$$

The Nyquist diagram of Eq. 18-98 as presented in Fig. 18-46 best illustrates the behavior of this sum.

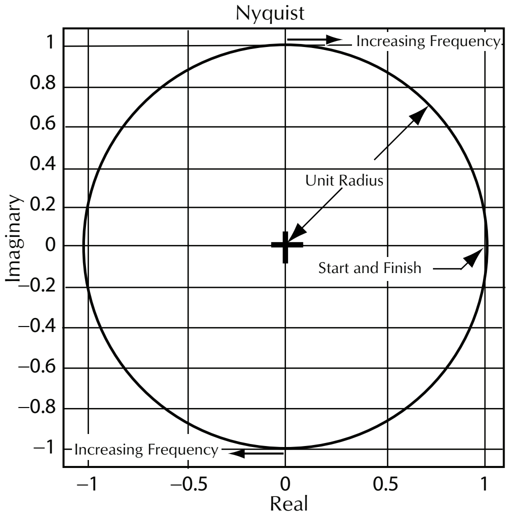


Figure 18-46. Nyquist diagram of summed frequency response of third order Butterworth.

The fact that the Nyquist diagram of Fig. 18-46 is a circle of unit radius centered on the origin means that the summed frequency response has a magnitude of unity or 0 dB at all frequencies and hence an all pass character. At zero frequency, the phasor representing the summed response is a horizontal line extending from the origin at (0,0) to the point (1,0*j*). The sum at zero frequency thus has a magnitude of one and an angle of zero. At each successive increasing value of frequency, the sum phasor maintains a constant length while having been rotated through increasing angles in the clockwise direction. When infinite frequency has been reached, the phasor will have rotated through one complete revolution. Exactly one-half of this rotation is accomplished as ω increases from zero to ω_0. This means that when $\omega = \omega_0$, the sum phasor extends from the origin to the point $(-1, j0)$. The remainder of the complete rotation is accomplished as the angular frequency increases from ω_0 to an infinite value.

The properties of the Butterworth polynomials are such that regardless of the order, the poles of the transfer functions always lie on a semi-circle of radius ω_0 in the left half of the complex plane. The poles are uniformly spaced with regard to angle with odd orders having a single pole on the negative real axis. Additionally, the zeros for the high pass function are all located at the origin. The consequence of this arrangement of poles and zeros is as follows.

The low pass function starts with a phase shift of zero and tends toward a phase shift of $-n\pi/2$ where n is the network order, having accomplished one-half of this total phase shift at the crossover frequency.

The high pass section starts with a phase shift of $+n\pi/2$ and tends toward zero as the frequency increases having accomplished a change of one-half of the total at the crossover frequency. Furthermore, the magnitude of the functions is equal at crossover with each having a value of $1/\sqrt{2}$ or −3 dB. Fig. 18-47 is a phasor diagram, at crossover, expressing these results for all orders through the seventh. The phasors in each case represent the signals passed by both the low and high frequency sections at crossover as well as the phasor sum.

A study of Fig. 18-47 reveals that even orders never sum to the appropriate value at the crossover frequency. The odd orders always sum properly at crossover and indeed at all frequencies even though the phase may or may not be correct. It was previously observed that the first order filters sum correctly in both magnitude and phase at all frequencies. It is unique in this respect. The present popularity of the higher odd order Butterworths is attributable to the fact that they sum to the appropriate magnitude at all frequencies including the crossover frequency. The transfer function describing this sum is of an all pass character but with a frequency dependent phase shift.

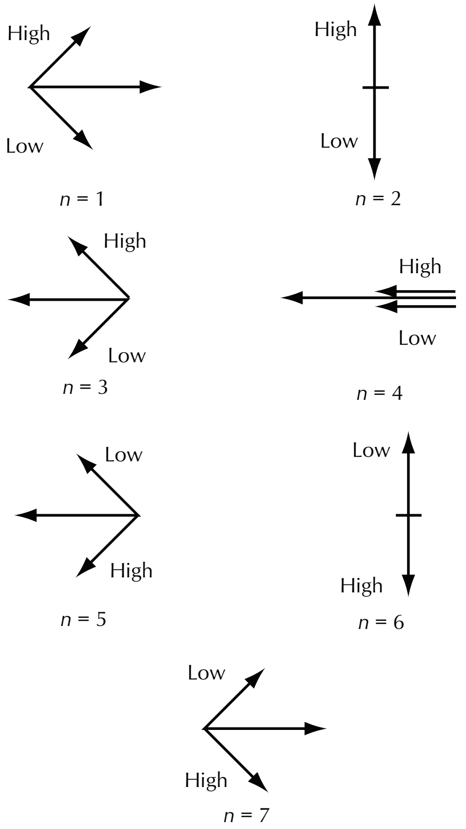


Figure 18-47. Butterworth amplitude summation properties for various orders.

All of the foregoing statements with regard to the summation properties of two-way crossover networks were made in reference to the summation of the electrical signals. In practice, however, it is the acoustical summation of the signals radiated by the low and high frequency loudspeakers which is of interest. In addition to correct crossover network behavior, the correct acoustical summation imposes additional conditions that must be satisfied by the loudspeakers employed.

In order to have correct power summation the loudspeakers must have reasonably flat amplitude and phase response in their respective pass bands. The loudspeakers must have equal axial Q and adjusted sensitivity at least for an octave both above and below crossover. Finally, proper amplitude response summation requires that the high and low frequency loudspeakers have a common acoustic center and acoustic origins which have been adjusted by signal delay techniques to be the same. This latter condition can be satisfied exactly only by co-axial loudspeakers. Other physical arrangements greatly restrict the region in the listening space in which the summation will be correct.

There are other possibilities for higher order crossover filters such as the Bessel or Chebyshev. Mathematically, these filters are distinguished by different sets of coefficients in the various terms of the polynomials in S in the filter transfer functions. The Bessel filters offer maximally flat group delay. This means that the phase response of this filter type is linear in the filter's pass band. The Chebyshev filters offer maximum attenuation rate at the edge of the filter stop band accompanied by ripple in the amplitude response in the pass band. Neither of these filter types is constant resistance in the passive implementation. Additionally these filters are lacking in unity summation properties. Their employment in crossover networks is usually driven by considerations other than those mentioned.

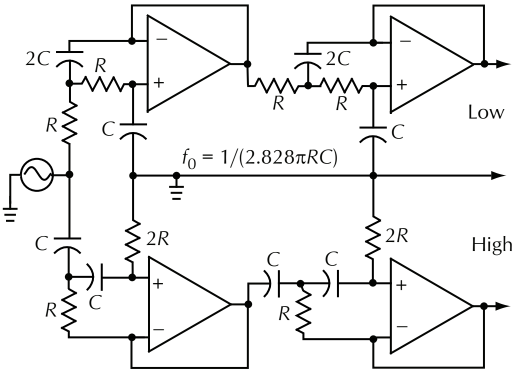


Figure 18-48. Linkwitz-Riley fourth order network.

There is one other two-way crossover network which is of more than just passing interest. This is the Linkwitz-Riley or second order Butterworth squared network. This network is almost always implemented as an active network by cascading two second order Butterworth networks as shown in the circuit of Fig. 18-48.

This network is fourth order, but it is not a fourth order Butterworth as it has a different set of polynomial coefficients. The low pass section has the transfer function of Eq. 18-99.

$$H_l = \frac{\omega_0^4}{S^4 + 2\sqrt{2}S^3\omega_0 + 4S^2\omega_0^2 + 2\sqrt{2}S\omega_0^3 + \omega_0^4} \tag{18-99}$$

The transfer function for the high pass section appears in Eq. 18-100.

$$H_h = \frac{S^4}{S^4 + 2\sqrt{2}S^3\omega_0 + 4S^2\omega_0^2 + 2\sqrt{2}S\omega_0^3 + \omega_0^4} \tag{18-100}$$

The behavior of the Linkwitz-Riley network is summarized in Figs. 18-49, 18-50, and 18-51. From Fig. 18-49 it is apparent that this network offers attenuation rates of 24 dB/octave and at the crossover frequency the response of each section is down by 6 dB rather than 3 dB as would be true for all orders of the Butterworth networks. Fig. 18-50 indicates that at the crossover frequency, the low frequency signal is in phase with the high frequency signal as +180° and −180° locate the same point on a unit circle. This implies that at the crossover frequency the total amplitude response is at 0 dB.

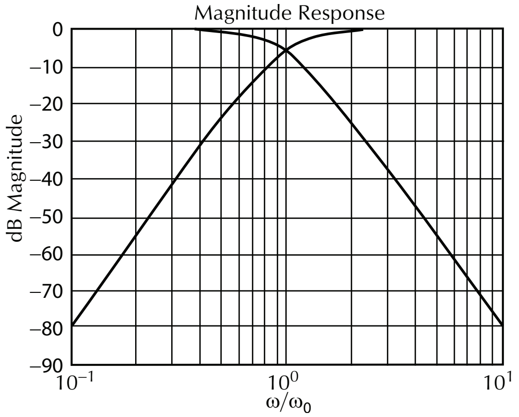


Figure 18-49. Linkwitz-Riley amplitude response.

Finally from Fig. 18-51, the Nyquist diagram is a circle of unit radius thus indicating that the summed amplitude response is 0 dB not only just at crossover but at all frequencies. Therefore the network is of an all pass character with an overall phase shift of 2π radians between frequency extremes. The shortcoming of this network lies only in its lack of unity power summation. In spite of this one shortcoming, the Linkwitz-Riley is often employed because of its structural simplicity and its steep slopes.

Three-Way and Higher Crossover Networks

The traditional approach for constructing a three-way crossover has been to employ a low pass at a low frequency and a high pass at a considerably higher frequency. These in turn border a bandpass whose lower half power point corresponds to that of the low pass and whose upper half power point corresponds to that of the high pass. A circuit for a simple version of such a network is illustrated in Fig. 18-52.

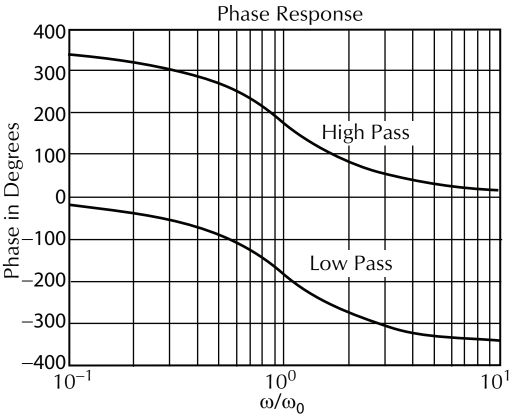


Figure 18-50. Linkwitz-Riley phase response.

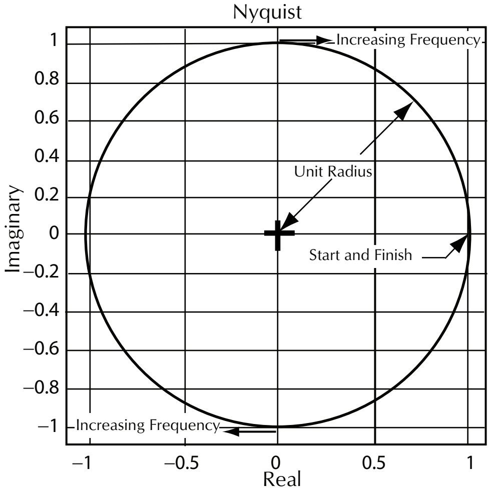


Figure 18-51. Linkwitz-Riley Nyquist diagram.

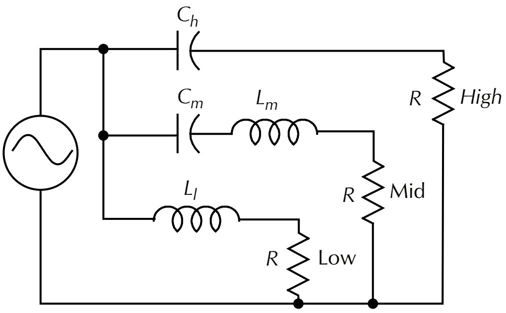


Figure 18-52. Simple passive three-way crossover circuit.

The transfer functions for the network of Fig. 18-52 are given in the following equations.

$$H_l = \frac{R}{L_l S + R} = \frac{\frac{R}{L_l}}{S + \frac{R}{L_l}} = \frac{\omega_l}{S + \omega_l} \tag{18-101}$$

$$H_m = \frac{R}{\frac{1}{C_m S} + L_m S + R} = \frac{\frac{R}{L_m} S}{S^2 + \frac{R}{L_m} S + \frac{1}{L_m C_m}} = \frac{\frac{\omega_m}{Q} S}{S^2 + \frac{\omega_m}{Q} S + \omega_m{}^2} \tag{18-102}$$

$$H_h = \frac{R}{\frac{1}{C_h S} + R} = \frac{S}{S + \frac{1}{RC_h}} = \frac{S}{S + \omega_h} \tag{18-103}$$

In Eq. 18-101, ω_l is the angular frequency of the half power point of the low pass filter. A specification of this value allows one to determine the required value of the inductance in this filter as, also from Eq. 18-101,

$$L_l = \frac{R}{\omega_l} \tag{18-104}$$

Similarly, in Eq. 18-103, ω_h is the angular frequency of the half power point of the high pass filter. A specification of this value allows one to determine the required value of the capacitance in this filter.

$$C_h = \frac{1}{\omega_h R} \tag{18-105}$$

The situation with regard to the bandpass or mid band filter described by Eq. 18-102 is a little more involved. Such filters are geometrically symmetric which is equivalent to the requirement that

$$\omega_m = \sqrt{\omega_h \omega_l} \tag{18-106}$$

Additionally, the Q or quality factor of this filter can be expressed as

$$Q = \frac{\omega_m}{\omega_h - \omega_l} = \frac{\omega_m L_m}{R} = \frac{1}{\omega_m C_m R} \tag{18-107}$$

Equations 18-106 and 18-107 may be solved simultaneously to obtain separate expressions for the required inductance and capacitance necessary for the construction of this filter.

$$L_m = \frac{R}{\omega_h - \omega_l} \tag{18-108}$$

$$C_m = \frac{\omega_h - \omega_l}{\omega_m{}^2 R} \tag{18-109}$$

The choice of crossover frequencies to be employed in three-way systems is driven by both the properties of the available drivers and acoustical considerations. From an acoustical point of view, it is highly desirable to make the reinforcement of speech as natural as is possible. This is greatly facilitated by having the mid range element handle all of the range of frequencies in which the energy of speech is dominant. A reasonable choice for the operating range of the mid range element would then be 300–3000 Hz. Employing these crossover frequencies with eight-ohm drivers leads to the following component values.

$$L_l = \frac{R}{2\pi f_l} = \frac{8}{(2\pi)(300)} = 4.24\ \text{mH}$$

$$C_h = \frac{1}{2\pi f_h R} = \frac{1}{(2\pi)(3000)8} = 6.63\ \mu F$$

$$\begin{aligned} L_m &= \frac{R}{\omega_h - \omega_l} \\ &= \frac{8}{(2\pi)(3000-300)} \\ &= 0.472\text{ mH} \end{aligned}$$

$$\begin{aligned} C_m &= \frac{\omega_h - \omega_l}{\omega_m{}^2 R} \\ &= \frac{(2\pi)(3000-300)}{(2\pi)^2(3000)(300)(8)} \\ &= 59.7\,\mu F \end{aligned}$$

The performance properties of a three-way network constructed with the above component values are best illustrated in a sequence of figures. Fig. 18-53 displays the amplitude performance of the simple three-way network.

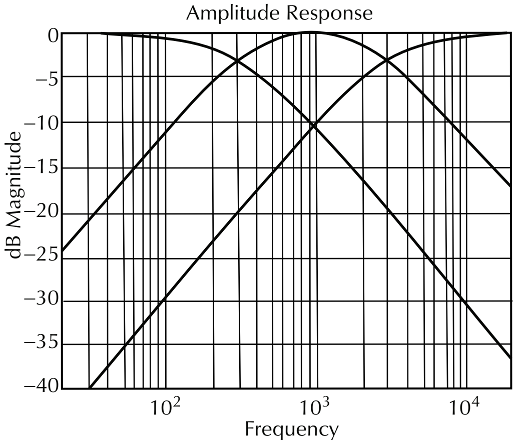


Figure 18-53. Amplitude response of a simple three-way crossover network.

From Fig. 18-53 it is apparent that the low to mid transition occurs at 300 Hz while the mid to high transition occurs at 3000 Hz. Additionally the bandpass center frequency occurs at the geometric mean of 300 Hz and 3000 Hz which is 949 Hz. The filter skirts have slopes of 6 dB/octave in each instance. The behavior of the summed frequency response appears in Fig. 18-54.

The summed frequency response of the simple three-way network is found to be not quite ideal. The ideal sum would be unity corresponding to zero dB at all frequencies. In practical terms, the worst case deviation is only a little over 1.4 dB and is not severe. Similarly, Fig. 18-55 displays the summed power behavior.

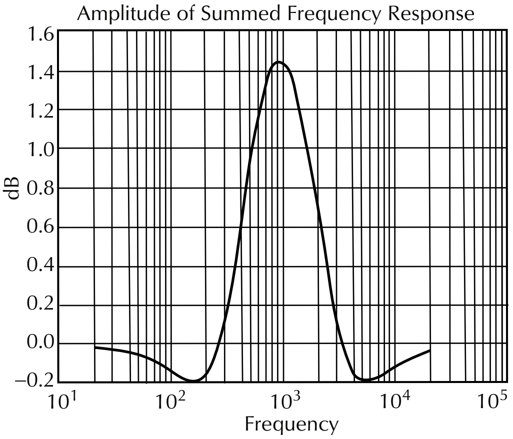


Figure 18-54. Simple three-way summed frequency response behavior.

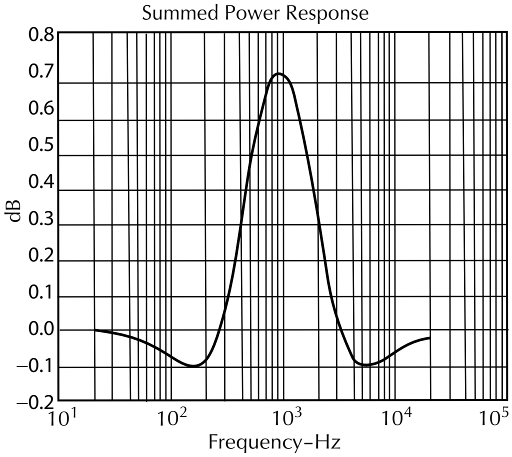


Figure 18-55. Simple three-way summed power response behavior.

Ideally the summed power response would be unity corresponding to 0 dB. The departure from ideal of a little over 0.7 dB displayed by the simple three-way is actually trivial in practical terms. The major shortcoming of this simple three-way lies in the low attenuation rate.

There are artful techniques for improving the performance of multi-way crossover networks. In fact, the simple three-way network can be greatly improved through a simple circuit change. As an example, consider the circuit of Fig. 18-56.

The low pass transfer function is the same as in the simple three-way.

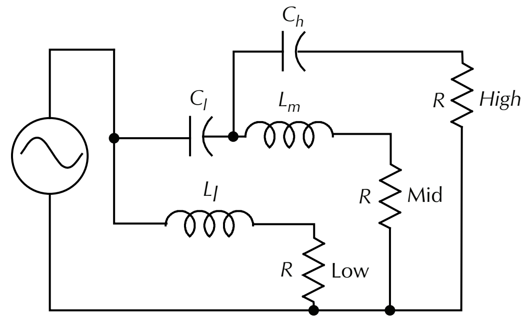


Figure 18-56. Improved three-way network.

$$H_l = \frac{\frac{R}{L_l}}{S + \frac{R}{L_l}} = \frac{\omega_l}{S + \omega_l} \tag{18-110}$$

The bandpass transfer function is now

$$\frac{S\frac{R}{L_m}}{S^2 + S\left(\frac{1}{RC_l} + \frac{R}{L_m}\right) + \frac{1}{L_m C_l}} \tag{18-111}$$

and the high pass transfer function becomes

$$H_h = \frac{S^2}{S^2 + S\left(\frac{1}{RC_l} + \frac{1}{RC_h}\right) + \frac{1}{R^2 C_l C_h}} \tag{18-112}$$

Upon taking 300 Hz and 3000 Hz as the transition points, the component values are found to be

$$L_l = \frac{R}{2\pi f_l} = \frac{8}{(2\pi)(300)} = 4.24\ \text{mH}$$

$$C_l = \frac{1}{2\pi f_l R} = \frac{1}{(2\pi)(300)8} = 66.3\ \mu\text{F}$$

$$L_m = \frac{R}{2\pi f_h} = 0.424\ \text{mH}$$

$$C_h = \frac{1}{2\pi f_h R} = \frac{1}{(2\pi)(3000)8} = 6.63\ \mu\text{F}$$

This rather modest change in the circuit arrangement and, as it turns out, a slight modification of two component values brings about marked improvement of performance in several areas. The improved network is a true constant resistance network. This was not the case for the original version. The summed frequency response is ideal in that the sum of the transfer functions is identically equal to one. The summed power response is also identically equal to one. Last, but by no means least, the slope of the high frequency filter is now 12 dB/octave rather than 6 dB/octave as is true for the former case. This increased slope of the high frequency filter offers more protection to the high frequency driver. The performance curves appear in the Figs. 18-57, 18-58, and 18-59.

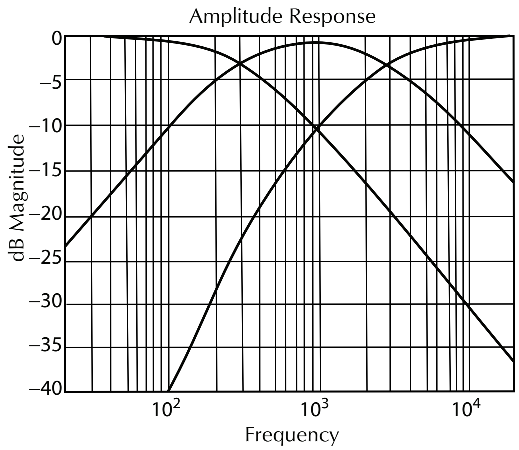


Figure 18-57. Amplitude response of improved three-way.

Synthesized Crossover Networks

Most current installed sound reinforcement systems employ active crossover networks with subsequent power amplification. Though more costly, this method offers a number of technical advantages. Differences in speaker sensitivities are readily compensated by power amplifier level adjustments. All speakers are fed from the low impedance sources offered by the power amplifier outputs.

Additionally, the active crossover systems whether analog or digital offer greater flexibility in

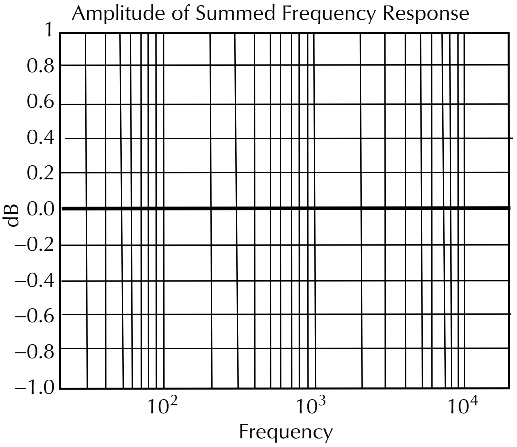


Figure 18-58. Summed frequency response behavior of improved three-way.

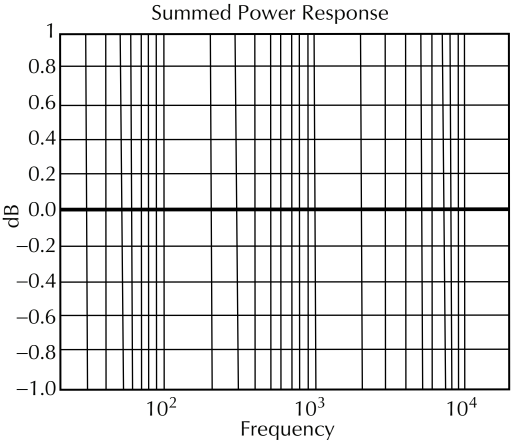


Figure 18-59. Summed power behavior of improved three-way.

generating filter transfer functions and, in the case of digital systems, variable signal delays for adjusting acoustic origins are readily available. The greater flexibility in generating filter transfer functions is manifested through the fact that in active systems, one may readily add or subtract signals. As an example, suppose one is working with a two-way system in which the high frequency drivers require at least a second order high pass filter in order to have adequate protection from low frequency signals.

Additionally, it is required that the low pass be such that the low and high frequency signals sum in such a way that the sum exhibits an all pass character with zero phase shift. This can be accomplished very simply in an active system. Consider the block diagram of Fig. 18-60.

In Fig. 18-60 the full audio spectrum signal is applied to both a conventional high pass filter of choice and to the non-inverting terminal of a unity gain difference amplifier. The output of the high pass filter is supplied as a separate output and is also applied to the inverting terminal of the unity gain difference amplifier. The signal passed by the high pass filter is thus subtracted from the total spectrum with the difference being constituted only of low frequency signals. In practice, the high pass could be any order desired even though for the present example a second order will be employed. In the language of transfer functions, the low pass transfer function generated by this arrangement will be simply

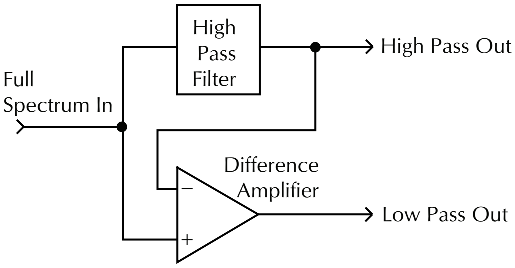


Figure 18-60. Synthesized low pass filter.

$$H_l = 1 - H_h \tag{18-113}$$

This procedure forces the sum of the high pass and low pass to be the ideal value of unity because

$$\begin{aligned} H_l + H_h &= 1 - H_h + H_h \\ &= 1 \end{aligned} \tag{18-114}$$

In particular, if the high pass filter selected is a second order Butterworth, the high pass transfer function is

$$H_h = \frac{S^2}{S^2 + \sqrt{2}S\omega_h + \omega_h{}^2} \tag{18-115}$$

while the synthesized low pass becomes

$$H_l = \frac{\sqrt{2}S\omega_h + \omega_h{}^2}{S^2 + \sqrt{2}S\omega_h + \omega_h{}^2} \tag{18-116}$$

Fig. 18-61 displays the amplitude responses associated with the network where 500 Hz has been chosen as the half power point for the high frequency section.

This technique, however, does have its own shortcomings. The low pass function necessary to guarantee unity summation in frequency response behavior is not maximally flat but rather exhibits a 2 dB peak. The attenuation slope of the low pass

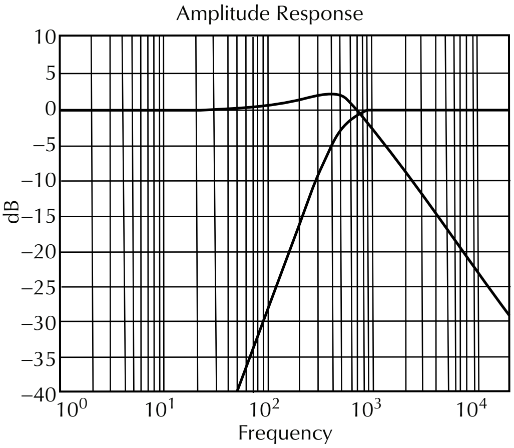


Figure 18-61. Amplitude response of two-way crossover with synthesized low pass.

section is only 6 dB/octave and the low frequency speaker must operate well at frequencies well beyond 500 Hz. Finally, even though the amplitude summation is unity as is also displayed in the figure, the power summation is not. The power behavior appears in Fig. 18-62.

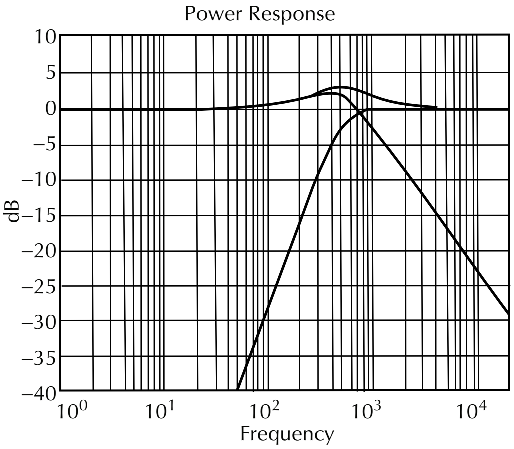


Figure 18-62. Power behavior of two-way with synthesized low pass filter.

Fig. 18-62 displays the relative power curves in the synthesized low pass section, the high pass section, and the overall summed value. The summed power curve exhibits a bump of about 3 dB that is reasonably localized and, though not ideal, is not all that severe.

18.13 Loudspeaker Arrays

Loudspeakers and loudspeaker systems are arrayed in order to produce coverage patterns that are unattainable from a single loudspeaker or loudspeaker system acting alone. Arraying, depending upon the techniques involved, can increase or decrease the total coverage angles. Arrays of two or more loudspeakers can be employed to increase the attainable sound pressure level.

Certain arrays of loudspeakers can modify the normal attenuation rate with distance at least in the intermediate field and open the possibility of electrically steering the directional pattern of the array. In former times arrays were built up of separate collections of low frequency and high frequency loudspeaker units. The low frequency units being either direct radiators, vented enclosures, or horn loaded. The high frequency units were almost universally horn loaded.

Such systems were obviously of the two-way variety. The trend in recent times has been to construct arrays from individual complete two or three way loudspeaker systems. The better of these systems are horn loaded with co-axial and or co-entrant arrangements between elements. The electrical drive signals applied to the low, mid, and high frequency elements are appropriately delayed to produce a common acoustic origin for all elements. In the best designs, the elements also share a common acoustic center. As a result, each such loudspeaker system can be considered as a single device for the purpose of constructing arrays.

Acoustic Order of Choices of System Structure

An array of loudspeaker systems is not always necessary. In a given environment, if the required level, coverage, and intelligibility can be obtained from a single, suitably installed loudspeaker nothing further is required or desired. In such an instance, particularly with regard to speech reinforcement systems, the only other consideration is source identification.

Source identification is the technique of locating the source of reinforced sound near the area where the original sound is produced. A case in point would be that of a live talker at a podium. If the loudspeaker is elevated above the podium by a few meters and pointed downward into the audience, the visual cues of the speaker's mouth movements and the audible cues of the natural and reinforced sound will fuse to give the overall impression that the total sound is emanating from the speaker's mouth.

In contrast, visualize the situation where you are facing the talker and the reinforced sound is coming from behind you. This is a situation not found in nature and leads to both confusion and fatigue. In addition, the downward projection of the loudspeaker into the audience leads to a useful value for the acoustic modifier, M_a. This in itself lowers the power driving the reverberant field and improves intelligibility. Similar such considerations lead to the following order of system choices to be explored in the design phase.

The system to be selected is the first one from this listing which satisfies the level, coverage, and intelligibility requirements for the acoustic space at hand.

1. A single source loudspeaker system.
2. An array of single source devices all at the same location.
3. A single source system accompanied by delayed satellite systems with each covering a distinct audience area.
4. High-density overhead distributed system.
5. Pew-back-type individual coverage.
6. Headphones for each listener.

Design Process

The prime considerations in system choice are those of coverage and intelligibility. One begins the design process by assuming a single source loudspeaker system and then determines what the system Q must be in order to satisfy intelligibility requirements for the most distant listener in the acoustical space at hand. If the space under consideration is an existing space, it is possible to make direct measurement of the reverberation times in octave or one-third octave bands. When the space exists only on architectural drawings one must rely on calculated values for reverberation times. In this instance a premium is placed on an accurate description of surface materials and treatments. In either case one takes the conservative approach and treats the space as if it is unoccupied. A fully occupied space has significantly lower reverberation times, which in turn produce higher intelligibility. If a system meets intelligibility requirements in an unoccupied space, its intelligibility only improves as the space is filled. The only exception to this rule is if the crowd noise worsens the signal to noise ratio such that it falls below 25 dB.

Consider a large multipurpose space that has a volume of 500,000 ft³ and a reverberation time of 4.3 s in the octave band centered on 2 kHz. An area devoted to lectures and other oral presentation is contained within this larger space. Fig. 18-63 is an elevation view through the centerline of the lecture space.

The relevant equation for determining the required Q for a single loudspeaker covering a distant listener is

$$Q_{MIN} = \frac{656(D)^2(RT_{60})^2}{(\%AL_{CONS})(V)} \tag{18-117}$$

where,

Q_{MIN} is the minimum required Q in the direction of the listener,

656 is a constant appropriate for English units,

D is the distance to the listener in feet,

RT_{60} is the reverberation time in seconds,

$\%AL_{CONS}$ is the percentage articulation loss of consonants,

V is the room volume in ft³.

When calculating with SI units, the constant 656 becomes 200, the distance is expressed in meters (m), and the volume is expressed in m³.

For the case depicted in Fig. 18-63, the most distant listener on axis is at a distance of 100 ft. The calculated minimum Q will then be the axial Q of the loudspeaker. From the equation then,

$$\begin{aligned} Q_{MIN} &= \frac{656(100)^2(4.3)^2}{(15)(500{,}000)} \\ &= 16.2 \end{aligned}$$

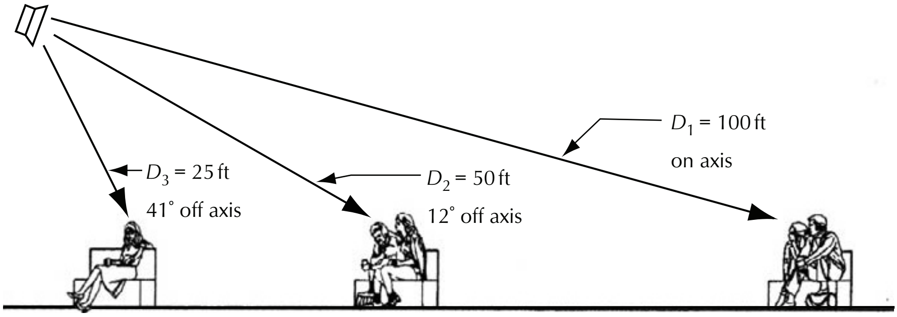


Figure 18-63. Elevation view.

As this is the minimum Q required, only an improvement will result by selecting a standard loudspeaker whose Q is slightly larger than the minimum. In this instance a loudspeaker whose nominal coverage is 60° by 40° with a Q on axis of 18 is a reasonable choice. The chosen loudspeaker has a sensitivity of 114 dB at a distance of 4 ft for a nominal electric power of 1 watt. The performance evaluation of this loudspeaker at the listening positions depicted in Fig. 18-62 consists of evaluating the direct sound levels at the respective seating positions and determining the $\%AL_{CONS}$ at these same positions.

The propagation distances to the various positions introduce a $1/r$ attenuation given by $20\ \text{dB}\log(4/D_i)$ where D_i corresponds to the respective listener distance. For those listeners not on the axis of the loudspeaker, there is an additional attenuation that must be obtained by consulting the polar response curves of the loudspeaker.

In this particular instance only the vertical polars are involved. The direct sound level at each position with one watt of excitation is then given by adding 114 dB to the total attenuation while remembering that the total attenuation is a negative value. The appropriate values are displayed in the following table.

Listener Position	Distance Attenuation	Polar Attenuation	Net Sound Level
Most Distant	–28.0 dB	0.0 dB	86.0 dB
Intermediate	–22.0 dB	–3.6 dB	88.5 dB
Nearest	–15.9 dB	–12.3 dB	85.8 dB

The difference between the highest and lowest sound level is less than 3 dB and is quite acceptable. Now it is necessary to determine the $\%AL_{CONS}$ at the various listener positions. In order to do this it is necessary to have values for the loudspeaker Q in the direction of the listener in question. The Q for the most distant listener whose location is on axis is just the axial Q value of 18 for the loudspeaker in question. For the intermediate listener as well as the nearest listener the appropriate Qs can be calculated by making use of the polar attenuation data.

The polar attenuation data tells one how much the sound intensity is reduced in the given angular direction. The governing equation is

$$Q_{\theta,\varphi} = Q_{AXIAL}10^{\frac{A}{10}} \qquad (18\text{-}118)$$

where,

$Q_{\theta,\varphi}$ is the Q in the specified direction,

Q_{AXIAL} is the axial Q,

A is the attenuation in the specified angular direction.

Once the various Qs have been calculated, one may determine the $\%AL_{CONS}$ from

$$\%AL_{CONS} = \frac{656(D)^2(RT_{60})^2}{(Q_{\theta,\varphi})(V)} \qquad (18\text{-}119)$$

The results for the case in hand appear in the following table.

Position	Polar Attenuation	Q Value	$\%AL_{CONS}$
On Axis	0.0 dB	18.00	13.5%
12° Below Axis	–3.6 dB	7.86	7.7%
41° Below Axis	–12.3 dB	1.06	14.3%

The maximum value of articulation loss is less than 15% thus indicating acceptable behavior. Based on the examination so far, the single loudspeaker solution appears to be satisfactory, as indeed it would be if all listeners were seated only on the centerline.

The seating pattern in this space is fan shaped being 50 ft wide at the rear tapering to 25 ft wide at the front. From both of the above tables, it appears that the nearest row will constitute the most severe problem. The outermost seat on the front row is 28 ft from the loudspeaker and thus has an increased distance attenuation of –16.9 dB. Furthermore this position is even further off axis of the loudspeaker such that the polar curves indicate an angle attenuation which has now become –15.8 dB. The sound level at this seat is thus reduced to (114 – 16.9 – 15.8) dB = 81.3 dB. The Q in the direction of this seating position from Eq. 18-118 now becomes only 0.47. The articulation loss at this seat as calculated from Eq. 18-119 is found to be 40%!!

Clearly, a single loudspeaker cannot adequately supply reinforcement in this space. This space is a candidate for a loudspeaker array type of solution. When faced with such a problem, the recommended practice is to employ a computer based loudspeaker design program. Not only will such programs save time in calculation, they also feature other important advantages.

An outstanding characteristic of such programs is the built-in loudspeaker database containing detailed information on polar response characteristics and other important operational parameters for commonly applied loudspeakers and loudspeaker systems. Additionally, provisions are made for updating the database or entering data manually. Many such programs feature extensive documentation capabilities for plotting both coverage and intelligibility patterns. Regardless of whether one is employing manual calculation or a computer based design program, it is necessary to have knowledge of successful loudspeaker array techniques.

Arraying Techniques

Arraying techniques are many and varied. Many of the early techniques evolved through trial and error. In more recent times the approach has become more increasingly based upon the physics of wave propagation. This became particularly true following the ready availability of computers capable of solving complicated wave propagation problems associated with multiple sources. In modern times, inexpensive digital signal processing techniques have made possible very versatile and finely tuned arrays whose characteristics can be modified under computer control.

Increase Coverage Angle

There are two well-recognized techniques for increasing coverage angle. These are stack and splay and trapezoidal array. The stack and splay can be applied to loudspeakers in general while the trapezoidal array is best applied to loudspeakers having an enclosure construction purposely designed for the application of this particular technique. Fig. 18-64 illustrates the stack and splay technique as applied to two alike constant directivity horns.

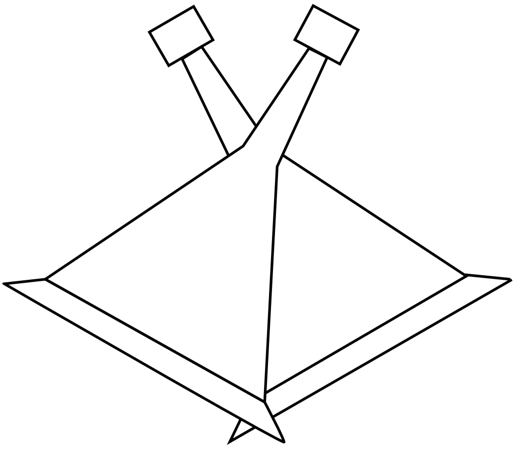


Figure 18-64. Stack and splay as applied to constant directivity horns.

Even though the illustration is drawn for horns, the technique is applicable to any loudspeaker system that possesses a well-defined acoustic center. An increase of horizontal coverage angle is obtained by splaying the loudspeakers about a vertical axis passing through the acoustic centers of both devices.

Alternatively, an increase of vertical coverage angle is obtained by splaying the loudspeakers about a horizontal axis passing through the acoustic centers of both loudspeakers. In both instances the angle of splay is the total angle between the central axis of each loudspeaker and is set equal to the common nominal coverage angle of each loudspeaker. This arrangement allows the radiation patterns of the individual loudspeakers to overlap at the half pressure points of the individual loudspeakers.

For the combination when so arrayed, the angle between the half pressure points of the combination will be just twice the nominal coverage angle of each loudspeaker in the plane of splay. To be specific, suppose one is splaying two devices each having a nominal coverage of 60° horizontal by 40° vertical with the objective of doubling the horizontal coverage angle to 120°.

One first stacks one device above the other followed by a rotation of one device clockwise through an angle 30° about a vertical axis passing through its acoustic center. Finally, the second device is rotated counter-clockwise through an angle of 30° about this same vertical axis. As observed in the far field, the horizontal coverage will now be 120° while the vertical coverage, though not as smooth as for an individual device, will approximate the original value of 40°.

The trapezoidal or trap array requires a loudspeaker system construction such that the acoustic centers of the devices in the plane of the array are located very near the rear of the loudspeaker system enclosures. The enclosures themselves have trapezoidal cross sections with the side angles being each equal to one-half of the coverage angle of the individual devices in the plane of the array. If the trap construction involves a constant directivity horn, the horn must be so oriented that its smaller coverage angle is in the plane of the array as the acoustic center for the smaller angle is the one closer to the rear of the horn. Fig. 18-65 depicts such an array of three trap elements each having a coverage angle of 40° horizontally by 60° vertically. The total coverage angle in the plane of such an array would be 120°.

With the arrangement of Fig. 18-65, the acoustic centers of the devices, though on a tight circular arc, are not coincident on the splay axis. In this arrangement, then, the combined coverage about the splay axis is reasonably smooth though not perfect whereas the vertical coverage is unperturbed from the original value.

Increase of Level on Axis

In those instances where insufficient acoustic pressure can be obtained from a single device, it is

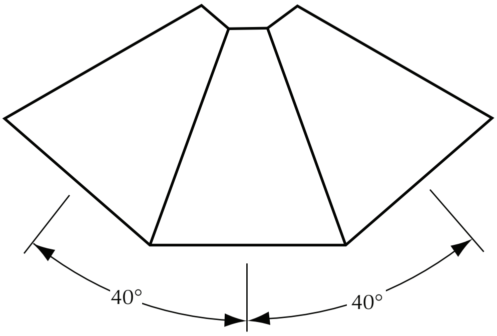


Figure 18-65. Three element trap array of 40° devices.

possible to stack two or more devices one above the other or to array two or more devices side by side. In both instances, the devices are alike and the device axes must be parallel. The physics and mathematics underlying this process will be explored in detail, as it will also be required in the study of line arrays. Fig. 18-66 depicts a stack of two devices though any number might be employed through simple extension.

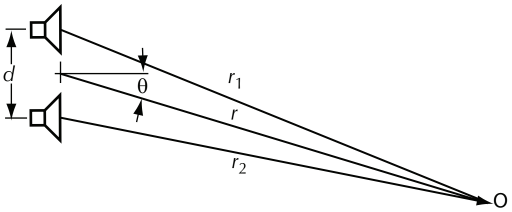


Figure 18-66. A stack of two devices for increasing acoustic pressure.

The objective is to determine an expression for the total acoustic pressure at the point O in terms of the variables r, θ, and the transducer properties. If the upper and lower transducers were simple sources of spherical waves, their individual acoustic pressures at O would be

$$
\begin{aligned}
p_1 &= \frac{A_1}{r_1} e^{j(\omega t - kr_1 - \varphi_1)} \\
p_2 &= \frac{A_2}{r_2} e^{j(\omega t - kr_2 - \varphi_2)}
\end{aligned}
\qquad (18\text{-}120)
$$

where,

A_1 and A_2 are the individual transducer amplitude factors,

φ_1 and φ_2 are arbitrary individual phase factors.

The total acoustic pressure at O is then the phasor sum of p_1 and p_2. When the objective is simply an increase of level, the individual amplitude factors are set equal and denoted by A while the individual phase factors are set equal to 0. Additionally, one examines the results in the far field where r is very much larger than d. When this is the case, it is possible to obtain a relatively simple expression for the total pressure through the following analysis. Upon applying the law of cosines to the two triangles of Fig. 18-66, the individual radial distances can be written exactly as

$$
\begin{aligned}
r_1^2 &= r^2 + \frac{d^2}{4} + rd\sin(\theta) \\
r_2^2 &= r^2 + \frac{d^2}{4} - rd\sin(\theta)
\end{aligned}
\qquad (18\text{-}121)
$$

Noting that r is very much larger than d, the above equations may be written with very little error as

$$
\begin{aligned}
r_1 &\approx r\sqrt{1 + \frac{d}{r}\sin(\theta)} \\
r_2 &\approx r\sqrt{1 - \frac{d}{r}\sin(\theta)}
\end{aligned}
\qquad (18\text{-}122)
$$

From the binomial theorem, it is known that the square root of one plus a quantity small compared with one is very nearly equal to one plus one-half of the small quantity. Therefore, Eq. 18-122 becomes

$$
\begin{aligned}
r_1 &\approx r + \frac{d}{2}\sin(\theta) \\
r_2 &\approx r - \frac{d}{2}\sin(\theta)
\end{aligned}
\qquad (18\text{-}123)
$$

If one sets the source amplitudes equal and sets the individual phase factors to zero, then substitution of Eq. 18-123 into Eq. 18-120 yields

$$
\begin{aligned}
p_1 &= \frac{A}{r + \frac{d}{2}\sin(\theta)} e^{j\left(\omega t - k\left[r + \frac{d}{2}\sin(\theta)\right]\right)} \\
p_2 &= \frac{A}{r - \frac{d}{2}\sin(\theta)} e^{j\left(\omega t - k\left[r - \frac{d}{2}\sin(\theta)\right]\right)}
\end{aligned}
\qquad (18\text{-}124)
$$

Now, in the denominators of Eq. 18-124 in each instance the second term is much smaller than the first and can be neglected with little error. Furthermore, the common terms in the exponential can be factored producing the result

$$p_1 = \frac{A}{r}e^{-jk\frac{d}{2}\sin(\theta)}e^{j(\omega t - kr)}$$
$$p_2 = \frac{A}{r}e^{jk\frac{d}{2}\sin(\theta)}e^{j(\omega t - kr)} \tag{18-125}$$

The total pressure is the phasor sum of p_1 and p_2 that can now be written as

$$p = \frac{A}{r}\left(e^{-jk\frac{d}{2}\sin(\theta)} + e^{jk\frac{d}{2}\sin(\theta)}\right)e^{j(\omega t - kr)} \tag{18-126}$$

One final simplification is now possible. Recall that from Euler's identity, $e^{jx} = \cos(x) + j\sin(x)$ and $e^{-jx} = \cos(x) - j\sin(x)$. Upon applying this to Eq. 18-126, the result appears as

$$p = \frac{2A}{r}\cos\left[k\frac{d}{2}\sin(\theta)\right]e^{j(\omega t - kr)} \tag{18-127}$$

The acoustic pressure of Eq. 18-127 is the product of three distinctly different factors. The first factor, $2A/r$, is essentially an amplitude term. The factor, $\cos[k(d/2)\sin(\theta)]$, is a directional modifier of the amplitude and as such is a directivity term. Lastly, the factor $e^{j(\omega t - kr)}$ is a complex exponential that describes wave propagation and has a magnitude that is always one. Information with regard to the pressure variation with regard to source strength, radial distance, direction, and frequency is all contained in the product of the first two factors. At the outset it was assumed that the two sources were simple spherical radiators with no directional characteristics of their own. In practice, each of the identical sources would have a directional characteristic in the plane of the array that in this instance is the vertical plane. As the devices are separated by only a small distance and observations are being made in the far field, the angle between either loudspeaker's principal axis and its own radial line is negligibly different from the angle θ. Upon denoting this common directional characteristic or vertical polar behavior as $f(\theta)$, the acoustic pressure amplitude of the two source arrays can now be written as

$$p_m = \frac{2A}{r}f(\theta)\cos\left[k\frac{d}{2}\sin(\theta)\right] \tag{18-128}$$

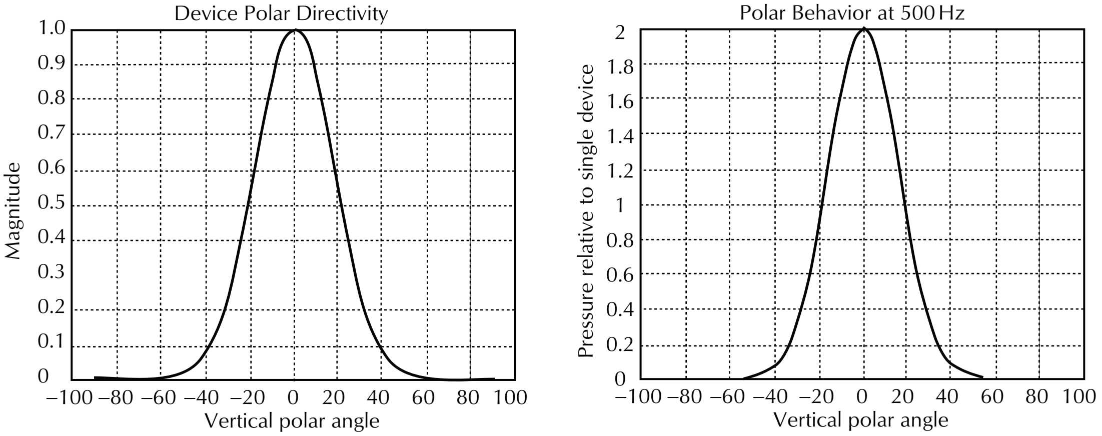
The pressure amplitude of Eq. 18-128 differs from the corresponding equation for a single device located at the origin in two significant ways. Firstly, the on axis pressure is twice as large and secondly, there is an additional directivity term that accounts for the interference between two sources that are not equidistant from the observer for all non-axial points.

At low frequencies, where k is small, and for small spacing between transducers, where d is small, this additional directivity term is nearly unity for all values of the polar angle. Above a critical frequency when the wavelength first becomes $2d$, this additional directivity term takes on more and more significance. As an example, consider two devices separated by a distance of 0.344 m (a little more than one foot) with each device having a half space vertical coverage angle of 40°. The critical frequency corresponding to this spacing is 500 Hz. Fig. 18-67A displays the vertical polar behavior of each device. Fig. 18-67B, C, and D display the array polar behavior at the critical frequency, at one octave, and two octaves above the critical frequency.

Below the critical frequency, the array vertical polar behavior is that of a single device from which the array is formed. At the critical frequency there is a small narrowing of the single lobe associated with the array vertical polar behavior. One octave above the critical frequency the central lobe is narrowed further and side lobes are first becoming evident. Two octaves above the critical frequency the side lobes have become significant and the central lobe is narrower still. Further increases in operating frequency bring about an exacerbation of this behavior with more numerous side lobes bordering an increasingly narrow central lobe. In this region, $f(\theta)$ serves only as an envelope of the multi-lobe pattern. A successful application of this technique at high frequencies will require high frequency devices that are physically small in order to achieve small device spacing thus making possible a large value for the critical frequency.

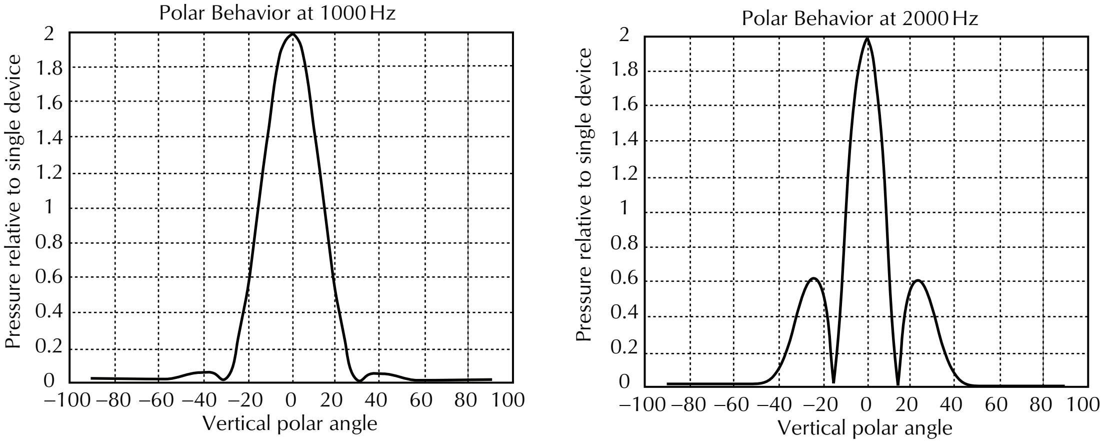
In the median horizontal plane, an observer is equidistant from each device and there is no departure from the normal horizontal directivity associated with each device. Many full range loudspeaker systems employ two bass, two mid range, and two high frequency devices. When this is the case, the bass loudspeakers are positioned at the extremities of the enclosure. The large separation being tolerated as the critical frequency is above the bass loudspeaker pass band.

The mid range speakers are sandwiched between the bass units and the high frequency units are sandwiched between the mid range devices. This arrangement allows medium spacing for the mid range and small spacing for high frequency devices thus allowing reasonable though not perfect operation of all devices in their respective pass bands.



A. Vertical polar behavior of each device.

B. Array vertical polar behavior at the critical frequency.



C. Array vertical polar behavior one octave above critical frequency.

D. Array vertical polar behavior two octaves above critical frequency.

Figure 18-67. Array vertical polar behavior.

18.14 Bessel Array

The Bessel array is another technique for increasing the attainable pressure level. This technique is not efficient from a power point of view but offers the distinct advantage that the array has a coverage pattern that is essentially identical to that of the individual loudspeakers constituting the array and without any frequency limitations. The successful operation of the array requires that all listeners truly be in the far field of the array. The far field requirement restricts Bessel arrays that are physically large (several feet in extent) to employment in large indoor arenas or in outdoor stadiums.

The Bessel array was first commercialized by Philips Corporation. The name of the array stems from the fact that the individual device amplitudes must have ratios mimicking those possessed by a sequence of Bessel functions. Bessel functions are solutions of Bessel's differential equation and were encountered previously with regard to the piston impedance and directivity functions. Bessel functions, when displayed graphically, have the general appearance of decaying sines or cosines though they are not periodic as are sinusoids. Bessel functions of various orders have been tabulated and can be generated by macros in most computer based math programs.

There are a variety of arrangements of devices for constituting a Bessel array. The most practical arrangement consists of five devices. The individual devices may in fact be full range loudspeakers but must be of the same type. The five devices may be arrayed either vertically or horizontally with the principal axes of the devices being exactly parallel. Fig. 18-68 depicts such a five element vertical array. The figure is not drawn to true scale for to do so would lose sight of speaker placement detail.

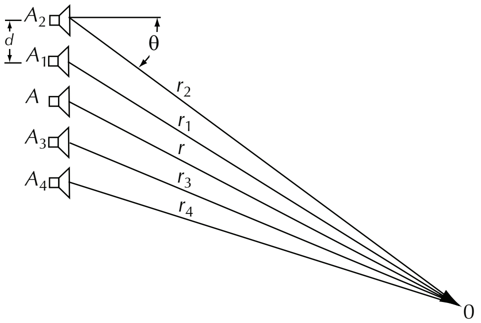


Figure 18-68. Five elements Bessel array.

The observer at point 0 is positioned in the far field where all of the radial distances in the figure are very large compared with the dimensions of the array. Furthermore, the angle between any radial line and the principal axis of its respective device departs from the value θ by a negligible amount. The directivity function common to each device is $f(\theta,\varphi)$. After applying the same type of analysis that was employed in treating the two element's array but now considering the angle to be negative as drawn, the total acoustic pressure at point 0 is found to be

$$p = \frac{f(\theta,\varphi)}{r}e^{j(\omega t-kr)}[A_4e^{-2jkd\sin\theta}+A_3e^{-jkd\sin\theta} + A + A_1e^{jkd\sin\theta}+A_2e^{2jkd\sin\theta}] \tag{18-129}$$

The As in Eq. 18-129 are amplitude coefficients that are proportional to the drive signals applied to the individual loudspeakers. The objective at this point is to find a magic assignment of amplitude coefficients that will make the term in the brackets equal to a constant or nearly so. This is the point at which Bessel functions enter the picture. One of the identity properties of Bessel functions can be expressed as

$$\left|\sum_{n=-\infty}^{n=\infty} J_n(x)e^{jn\alpha}\right| = 1 \tag{18-130}$$

In the above equation, x and α are any real independent variables. Eq. 18-130 requires a sum over an infinite number of terms in order for the absolute magnitude of the sum to take on an exact value of unity. A useful though not quite exact result can be obtained by employing only five terms from this infinite sum and placing them in correspondence to the sum of the five terms in Eq. 18-129. Upon setting α equal to ($kd\sin\theta$), the five terms from the sum of Eq. 18-130 are

$$J_{-2}(x)e^{-2jkd\sin\theta}+J_{-1}(x)e^{-jkd\sin\theta}+ J_0(x)+J_1(x)e^{jkd\sin\theta}+J_2(x)e^{2jkd\sin\theta} \tag{18-131}$$

The argument of the Bessel functions is the independent variable x with x allowed to take on any arbitrary real value. The task at this point is to find a value of x for which the ratios of the values of the various Bessel functions form a useful and convenient sequence. This may be accomplished by trial and error or more realistically by overlaying plots of the various functions versus x. Upon taking the latter approach it is found that if x is assigned the value of 1.5 then

$$J_{-2}(1.5) : J_{-1}(1.5) : J_0(1.5) : J_1(1.5) : J_2(1.5) \approx 0.5 : -1 : 1 : 1 : 0.5 \tag{18-132}$$

When this assignment is made in Eq. 18-129 the result can be written as

$$p = \frac{A}{r}f(\theta,\varphi)e^{j(\omega t-kr)}[0.5e^{-2jkd\sin\theta}-1e^{-jkd\sin\theta}+ 1+1e^{jkd\sin\theta}+0.5e^{2jkd\sin\theta}] \tag{18-133}$$

Now if one applies Euler's identity to the exponential terms in the bracket and collects terms, the pressure expression can be written as

$$p = \left(\frac{2A}{r}f(\theta,\varphi)e^{j(\omega t-kr)}\right)\times [\cos^2(kd\sin\theta)+j\sin(kd\sin\theta)] \tag{18-134}$$

The factor before the bracket in Eq. 18-134 describes a pressure wave of double strength emanating from a single loudspeaker having all of the usual directional and frequency characteristics of the radiating device. It remains only to explore the behavior of the bracketed expression. The expression in the brackets can be written as a complex exponential that expresses both the magnitude and phase behavior of the expression. Denoting this as a function D with D depending on d, k, and polar angle θ, the result is

$$\begin{aligned} D(d,k,\theta) &= Me^{j\Phi} \\ &= \sqrt{\cos^4(kd\sin\theta)+\sin^2(kd\sin\theta)} \\ &\quad\times\left(e^{ja\tan\left(\frac{\sin(kd\sin\theta)}{\cos^2(kd\sin\theta)}\right)}\right) \end{aligned} \tag{18-135}$$

One can readily evaluate both the magnitude and the phase of D for the on axis position where θ is zero. For $\theta = 0$, $\sin\theta = 0$ making the radical expression, and consequently M have the value unity at all frequencies. Remember that the frequency enters through the appearance of k as $k = \omega/c = 2\pi f/c$. Additionally, when $\theta = 0$, the angle Φ or phase of D is zero at all frequencies. In summary, for those listeners located on axis, D plays no role whatsoever. For observers located off axis, the magnitude of D will ripple between unity and 0.866 in a continuous fashion as the operating frequency increases.

This is of little concern as this variation amounts to only 1.25 dB peak to peak. Of perhaps more concern, though apparently undetected by most observers, is that the phase of D also ripples continuously between plus and minus $\pi/2$ or 90° as the operating frequency increases. The rapidity of this variation is more pronounced for larger device separation. This behavior is displayed in Figs. 18-69 and 18-70 where the device spacing is 0.6 m (about 2 ft) and the polar angle is fixed at 45°.

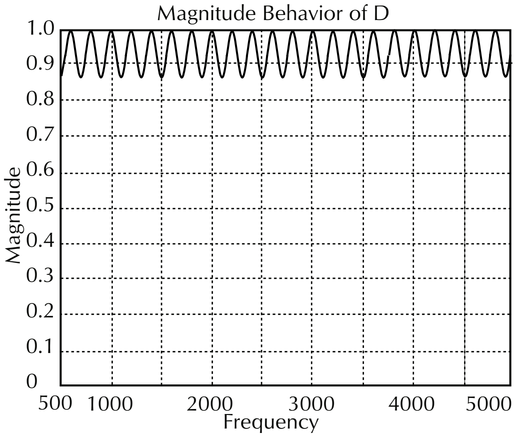


Figure 18-69. Ripple in Bessel array off-axis amplitude response.

The required voltage drive to the elements constituting the Bessel array of Fig. 18-68 may be accomplished by employing a single amplifier to drive all five loudspeakers or by using five individual amplifiers. Recall that A_4 and A_2 must present half value amplitudes. This can be accomplished by connecting these two devices in series before connecting the combination to a single amplifier or by using two amplifiers with each being attenuated by 6 dB. A_3 must present full value amplitude with negative polarity. The connections to this element then require just a reversal of polarity. A_1 and A must present full amplitude with normal polarity. A, A_1, and a reversed polarity A_3 can be placed in parallel with the series connected A_2 and A_4 with the entire combination being connected to a single power amplifier.

Alternatively, one can employ individual amplifiers for each element in the array. This would require two amplifiers attenuated by 6 dB with normal polarity, two unattenuated amplifiers with normal polarity, and one unattenuated amplifier with reversed polarity.

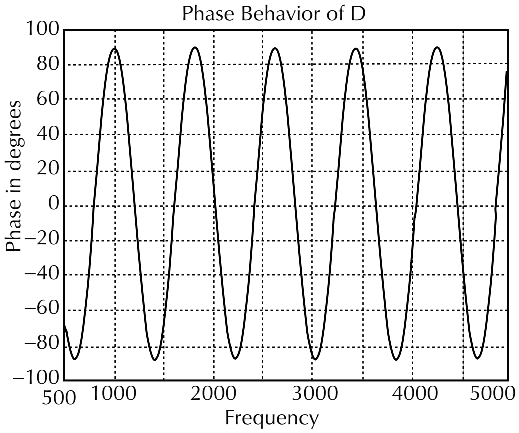


Figure 18-70. Ripple in Bessel array off-axis phase response.

18.15 Line Arrays

A true continuous line source is modeled as depicted in Fig. 18-71. The construction is that of an elongated cylinder of small radius wherein the radius alternatively expands and contracts by a small amount about its nominal value. If the cylinder were infinitely long, this would be a source of cylindrical waves for which the attenuation rate with perpendicular distance from the cylinder would be 3 dB for each doubling of the distance.

Even with a cylinder of finite length, when observations are made in the median plane close to the cylinder this same attenuation rate is found to be true. At larger distances, however, the attenuation rate transitions to the far field value of 6 dB for doubling of distance typical of any small source having in phase surface velocity. The distance at which the far field begins depends on both the length of the continuous line source and the operating frequency. A simple analysis provides an insight on where the far field begins. Consider the situation depicted in Fig. 18-72.

For the observation point depicted in the figure, radiation from the end of the source must travel a distance that is $\lambda/4$ greater than that traveled from

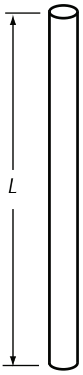


Figure 18-71. Continuous line source.

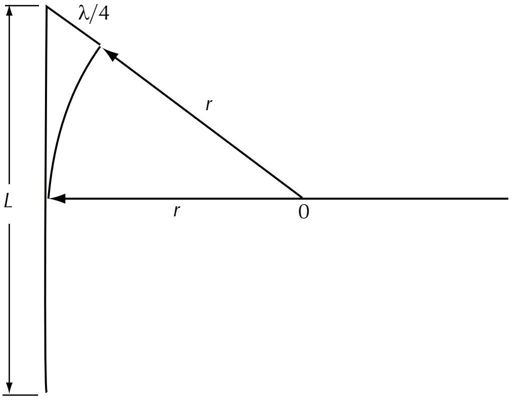


Figure 18-72. Far field calculation aid.

the center of the source. These two component radiations differ in phase by 90° and hence combine quadratically rather than linearly. This addition is still constructive as the phasor sum is greater than the individual component values. On the other hand, if the observation point is moved to a position slightly closer to the source thus shortening r, the arc will now intercept the line drawn between the end of the source and the new observation point at a distance that slightly exceeds $\lambda/4$. The phase difference between the two component radiations now will exceed 90° by a small amount and the phasor sum will be less. By the contrary argument, for observation points more remote than those indicated in the figure, the phase difference tends to zero as r is progressively increased. The present position of O defines the transition point between the near and far fields for the continuous line source. A relationship among the pertinent quantities is obtained by applying the Pythagorean Theorem to the triangle in the figure.

$$\left(\frac{L}{2}\right)^2 + r^2 = \left(r + \frac{\lambda}{4}\right)^2 \tag{18-136}$$

As will soon be learned, successful employment of line arrays requires that the array length is equal to or exceeds the wavelength throughout the operating frequency range. This allows the extraction from Eq. 18-136 of the following approximate statement of the radial distance at which the far field begins.

$$r \geq \frac{L^2}{2\lambda} \tag{18-137}$$

The length of the array, L, is dictated by the lowest operating frequency where the wavelength is the largest. The radial distance constituting far field operation, however, is proportional to the operating frequency, being quite large where the wavelength is the least. These results will be of importance later when continuous line sources are approximated by arrays of discrete loudspeaker elements.

The items of interest to the sound system practitioner are the acoustic pressure amplitude and the directivity characteristics of a line source built up with a discrete collection of loudspeakers. More importantly, perhaps, are answers to the questions relating to how to arrange a collection of loudspeakers so as to obtain a particular pressure amplitude and coverage pattern in a given direction. Interestingly, this is best pursued by studying continuous line sources and then modifying the results to accommodate a discrete source collection. Fig. 18-73 is the starting point for calculating the acoustic pressure produced in the far field by a continuous line source of finite length.

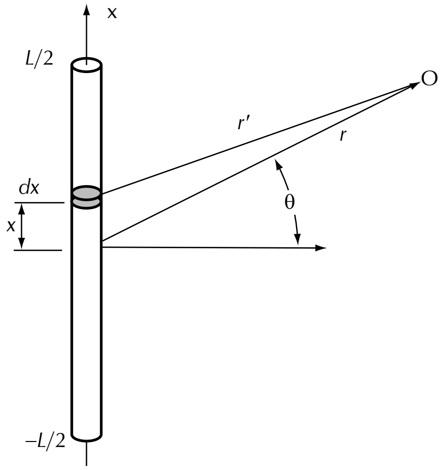


Figure 18-73. Continuous line source.

In Fig. 18-73 the shaded infinitesimal cylinder has a nominal radius a and a lateral surface area of $2\pi a dx$. This cylindrical element is alternately expanding and contracting a small amount about its nominal radius such that its surface velocity is represented by the phasor $u = u_m e^{j\omega t}$. The acceleration attendant to this motion is $j\omega u$. This element produces an infinitesimal acoustic pressure whose phasor description at the observation point O is given by

$$dp = \frac{\rho_0 j\omega u_m 2\pi a dx}{4\pi r'} e^{j(\omega t - kr')} \tag{18-138}$$

The total acoustic pressure at the observation point involves integrating the above expression over the entire length of the line. Recalling that $k = \omega/c$, the integral expression becomes

$$p(r', t) = \frac{j\rho_0 c u_m ka}{2} \int_{\frac{-L}{2}}^{\frac{L}{2}} \frac{e^{j(\omega t - kr')}dx}{r'} \tag{18-139}$$

The observation point is taken to be in the far field where $r' \approx r - x\sin\theta$. Upon making this substitution, the integral becomes

$$p(r, \theta, t) = \frac{j\rho_0 c u_m ka}{2r} e^{j(\omega t - kr)} \int_{\frac{-L}{2}}^{\frac{L}{2}} e^{jkx\sin\theta}dx \tag{18-140}$$

Note that in writing Eq. 18-140 r' in the denominator has been replaced simply by r. This is justified because in the far field the very small difference between r' and r makes an insignificant difference in the amplitude. On the other hand this small difference may be significant as compared with the wavelength and may cause an appreciable phase shift. This requires that the complete expression for r' be employed in the exponential expression. The integration can be readily performed to obtain the result

$$p(r, \theta, t) = \frac{j\rho_0 c u_m akL}{2r} \left(\frac{\sin\left(\frac{1}{2}kL\sin\theta\right)}{\frac{1}{2}kL\sin\theta} \right) e^{j(\omega t - kr)} \tag{18-141}$$

The significant results of the calculation are twofold. Firstly, the pressure amplitude for all far field points in the median plane where $\theta = 0$ is just the magnitude of the leading factor in the equation.

$$p_m(r) = \frac{\rho_0 c u_m akL}{2r} \tag{18-142}$$

Secondly, the radiation in the far field is also controlled by a directivity function $D(\theta)$ where

$$D(\theta) = \frac{\sin\alpha}{\alpha} = \frac{\sin\left(\frac{1}{2}kL\sin\theta\right)}{\frac{1}{2}kL\sin\theta} \tag{18-143}$$

The function $(\sin\alpha)/\alpha$ is an often encountered function known as sinc α. A plot of this function conveniently displays the general behavior of the directivity associated with a continuous line radiator. This plot is presented in Fig. 18-74.

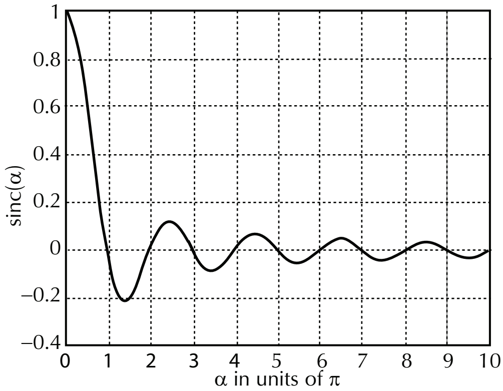


Figure 18-74. The behavior of the sinc function.

The more conventional manner of examining the directivity is by viewing polar plots of the absolute magnitude of the directivity versus the polar angle θ. This is done in the following sequence presented in Fig. 18-75 wherein the emphasis is on the relationship between the line length L and the wavelength of the radiation emitted.

The far field radiation from a continuous line radiator whose length is much shorter than the wavelength at its operating frequency is isotropic. This means that the radiation pattern in the far field is the same at all angles with no control of directivity. When the operating frequency is increased to the point that the radiator length is the same as the wavelength as depicted in Fig. 18-75A, directivity is well established with exactly just a single central lobe. The pattern in three dimensions would be obtained by rotating the figure about a line through the 90° and 270° positions. This line is the same as the axis of the continuous line radiator.

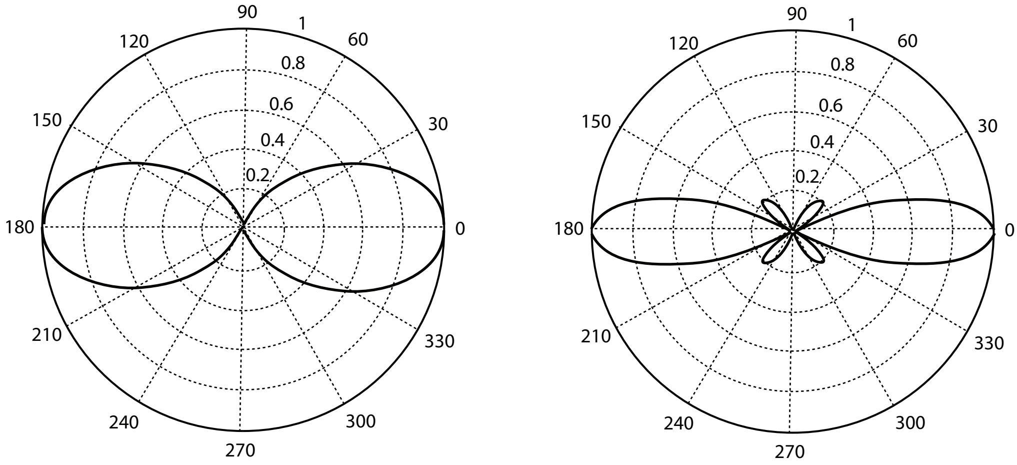
In Fig. 18-75B, the operating frequency has been doubled so that the line has become two wavelengths long. The pattern now contains a narrower central lobe bordered by two side lobes. This trend continues in Fig. 18-75C and 18-75D wherein raising the operating frequency further narrows the central lobe and introduces more side lobes.

Obviously, a uniformly excited continuous line source is not a constant directivity device. It could be made so, however, if its length could be altered such that it was inversely proportional to the operating frequency. This cannot be accomplished in a practical sense but it does suggest a possible alternative. Suppose it were possible to vary the drive amplitude as a function of position along the line. One out of many possibilities would be to linearly taper the drive amplitude in both directions from the center such that the drive amplitude falls uniformly from a maximum at the center to zero at $x = \pm L/2$.

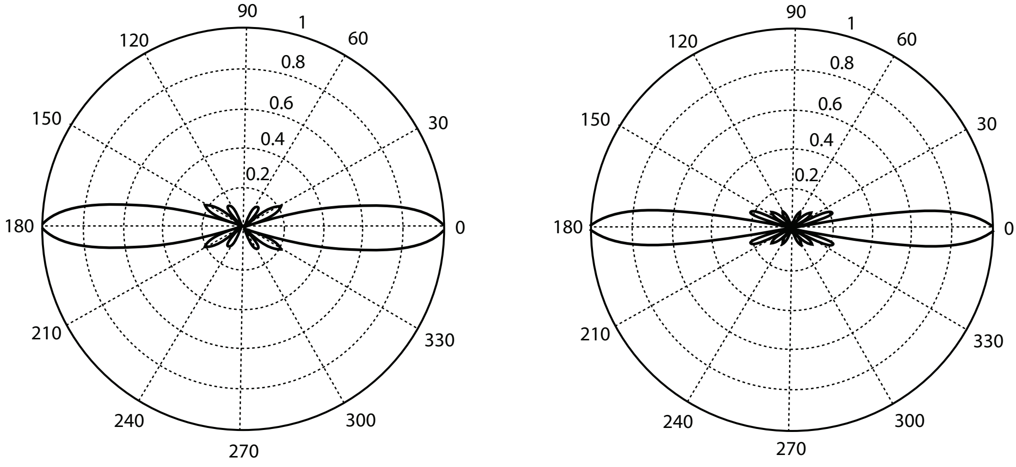
The calculation of Eq. 18-140 must now be repeated with the positional dependent amplitude function included under the integral. This results in a modified directivity function given by

$$D(\theta) = \left(\frac{\sin\left(\frac{1}{4} kL \sin\theta\right)}{\frac{1}{4} kL \sin\theta} \right)^2 \qquad (18\text{-}144)$$

The polar pattern of this new directivity function appears in Fig. 18-76. This figure is to be compared with Fig. 18-75B.



A. Line length = λ B. Line length = 2λ



C. Line length = 3λ D. Line length = 4λ

Figure 18-75. Continuous line polar patterns for various wavelengths.

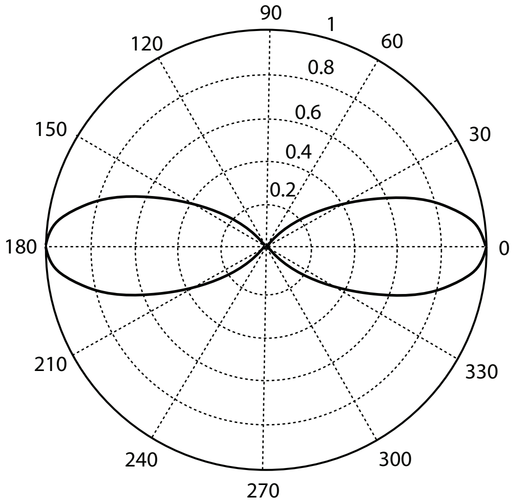


Figure 18-76. Amplitude tapered line of length 2λ.

In comparing the two figures it should be noted that linear amplitude tapering offers the dual benefits of greatly attenuating the side lobes while simultaneously widening the main lobe.

Line Array of Discrete Elements

Continuous line arrays are mostly theoretical constructs as the technology necessary to the construction of a truly continuous line array does not presently exist. The conclusions drawn from a study of continuous line arrays do provide useful guidance toward the construction of line arrays of discrete elements.

The study of the continuous line showed that the line length had to be comparable to the wavelength at the lowest operating frequency in order to obtain directional control. On the other hand, when the line is several wavelengths long, the central lobe becomes quite narrow with significant acoustic energy going into a multiplicity of side lobes. Furthermore, the study of the interaction of just two discrete devices concluded that the device spacing needed to be comparable to the wavelength at the highest frequency of operation.

These conclusions, when taken together, suggest that separate distinctly designed arrays be employed for low, mid, and high frequency portions of the spectrum. Fortunately the design equations are the same regardless of the frequency range to be covered. The number of discrete elements, size of element, and element spacing will vary, however, depending on the frequency range involved.

Fig. 18-77 depicts the spatial description of a generic discrete element line array. The total number of elements involved is arbitrary. It is assumed, however, that the elements have identical properties.

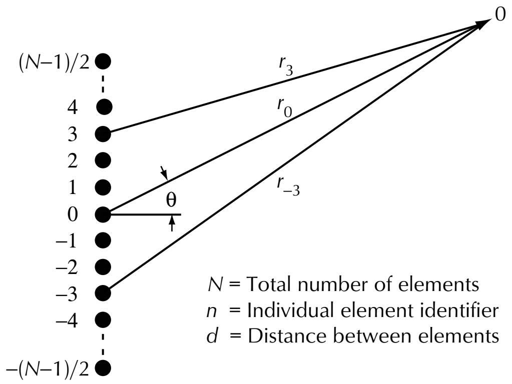

0
(N–1)/2
r_3
r_0
4
3
2
1
0
θ
r_{-3}
–1
–2
–3
–4
–(N–1)/2
N = Total number of elements
n = Individual element identifier
d = Distance between elements

Figure 18-77. Generic line array of *N* elements.

The black circles in Fig. 18-77 represent discrete loudspeakers that may be in individual enclosures or all mounted in a common enclosure. In either event all of the loudspeakers possess a common directivity function $D(\theta, \varphi)$ that is independent of the directivity function brought about by their physical arrangement in a line. Each of the loudspeakers may in fact be horn loaded for example. In the first approach to this problem it will be considered that the line is operating without any processing. This means that each source is driven with the same strength and all sources are driven in phase. The acoustic pressure produced at a far field observation point O located in the plane of the figure by any individual source is then

$$\begin{aligned} p_n &= \frac{A}{r_n} D(\theta, \varphi) e^{j(\omega t - kr_n)} \\ &\approx \frac{A}{r_0} D(\theta, \varphi) e^{j(\omega t - kr_0)} e^{jnkd\sin\theta} \end{aligned} \qquad (18\text{-}145)$$

In writing the second of Eq. 18-145 use has been made of the fact that $r_n \approx r_0 - nd\sin\theta$ and the usual far field substitutions were made. The total acoustic pressure at the observation point is the phasor sum of the individual contributions expressed as

$$p(r_0, t, \theta, \varphi) \approx \qquad (18\text{-}146)$$

$$\frac{A}{r_0} D(\theta, \varphi) e^{j(\omega t - kr_0)} \left(\sum_{n = \frac{-(N-1)}{2}}^{n = \frac{+(N-1)}{2}} e^{jnkd\sin\theta} \right)$$

Through the use of Euler's identity and some trigonometric identities the summation can be carried out to produce the result

$$p(r_0, t, \theta, \varphi) \approx \frac{A}{r_0} D(\theta, \varphi) e^{j(\omega t - kr_0)} N \left(\frac{\sin\left(\frac{N}{2} kd \sin\theta\right)}{N \sin\left(\frac{1}{2} kd \sin\theta\right)} \right) \tag{18-147}$$

The quantity of interest to most observers is the acoustic pressure amplitude as it depends on distance and direction. This can be extracted from Eq. 18-147 and can be written as

$$p_m(r_0, t, \theta, \varphi) = \left| N \frac{A}{r_0} D(\theta, \varphi) D_l(\theta) \right| \tag{18-148}$$

where,

$$D_l(\theta) \equiv \left(\frac{\sin\left(\frac{N}{2} kd \sin\theta\right)}{N \sin\left(\frac{1}{2} kd \sin\theta\right)} \right). \tag{18-149}$$

$D_l(\theta)$ is the directivity function of the line array itself independent of other factors. The two directivity factors $D(\theta, \varphi)$ and $D_l(\theta)$ can independently be positive or negative. This is the reason for the absolute magnitude symbol, $|\,|$, bracketing the right hand of Eq. 18-148 as the pressure amplitude is always positive. Additionally, the expression for $D_l(\theta)$ appearing in Eq. 18-149 is valid for any number of elements composing the array as long as the elements are identical, in phase, and of equal amplitude. The directional behavior of such an array is best viewed as a polar plot of the absolute magnitude of $D_l(\theta)$ versus the angle θ. Such plots are presented in Fig. 18-78.

In Fig. 18-78A through 18-78D the operating frequency is successively increased in octave steps. In Fig. 18-78A, the operating wavelength is twice the length of the array and there is little directional control. In Fig. 18-78B where there is equality between line length and operating wavelength, directional control is well established except for the small side lobes. The trend of a narrowing main lobe with increasing number of side lobes as the operating frequency continues to increase is evidenced in Fig. 18-78C and 18-78D.

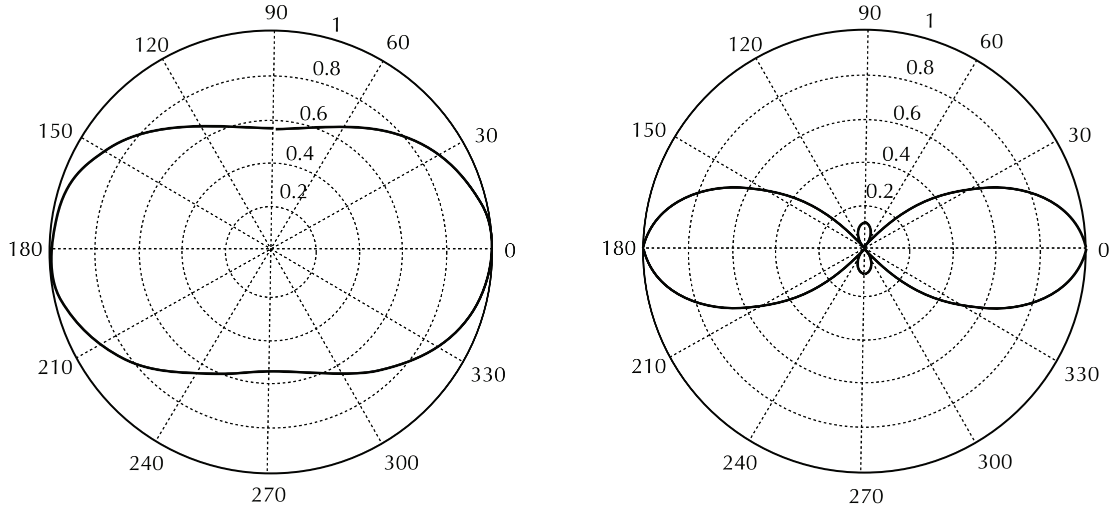
The desirable attribute of the line array to this point is simply that the maximum acoustic pressure amplitude in the main lobe is N times that produced by a single source. Fortunately, however, the ready availability of digital signal processing technology affords many tools for improving the behavior of discrete element line arrays. These features will be explored in the next section.

Processed Line Arrays

The availability of economical programmable digital signal processing circuitry as well as compact power amplifiers makes possible the complete tailoring of the drive signal applied to each element of an extended line array. The drive signals to the individual elements can be separately filtered, equalized, delayed, adjusted for level, and amplified before application to the individual transducers. This flexibility allows the shaping of the beam width of the radiation pattern of the array, the steering of the beam over a range of directions, focus of the beam at a particular point in space, and adjustment of the effective line length as a function of frequency. These properties will be explored through a sequence of examples beginning with beam steering. Fig. 18-79 depicts the geometrical situation in the immediate vicinity of an array of nine elements where it is desired that the radiation from the individual elements progresses in phase along a particular angular direction denoted by a steering angle θ_s.

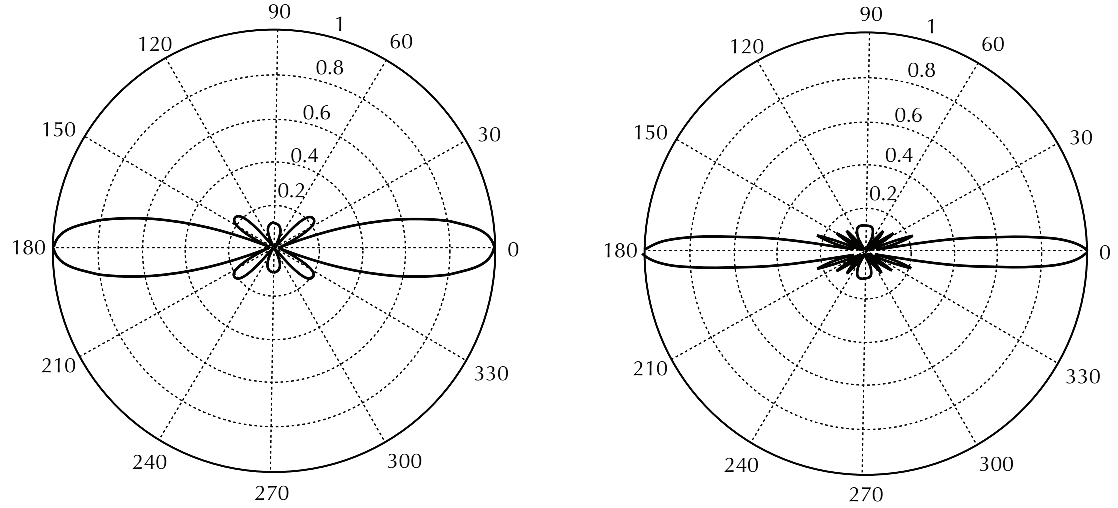
From the figure it is apparent the source denoted as 4 is nearest to the far field point. The source denoted as 3 is more distant by $d\sin\theta_s$, the source denoted as 2 is more distant by $2d\sin\theta_s$, etc., until one reaches the source denoted as –4 that is more distant by $8d\sin\theta_s$. The basic incremental radial distance between elements, Δr, is thus $d\sin\theta_s$. The basic incremental signal delay between elements is then $\Delta r/c$. In order for the signals to progress in phase in the chosen direction, the most distant element, –4, must be driven directly, –3 must have a signal delay of $\Delta r/c$, –2 must have a signal delay of $2\,\Delta r/c$, etc. The delay sequence continues in this fashion until one reaches source 4 where the required signal delay is $8\Delta r/c$. This procedure can be applied to an array of an arbitrary number of elements. When the appropriate delays are inserted in the summation of Eq. 18-146, the directivity function associated with the steered line alone becomes

$$D_l(\theta) \equiv \left(\frac{\sin\left[\frac{N}{2} kd(\sin\theta - \sin\theta_s)\right]}{N \sin\left[\frac{1}{2} kd(\sin\theta - \sin\theta_s)\right]} \right) \tag{18-150}$$



A. Line length = λ/2

B. Line length = λ



C. Line length = 2λ

D. Line length = 4λ

Figure 18-78. Eight element unprocessed line array polar behavior.

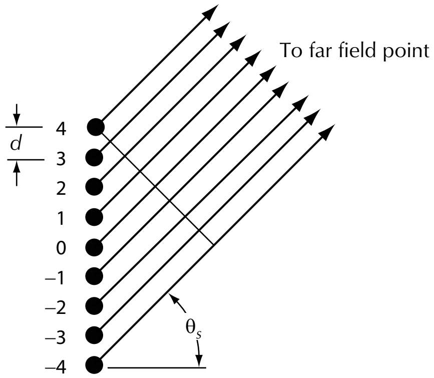
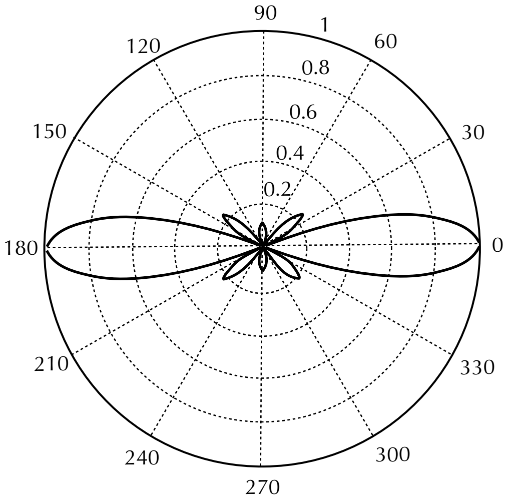


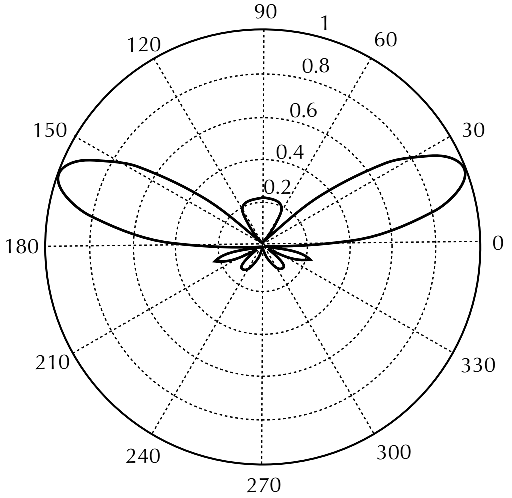
Figure 18-79. Steered array geometry.

Fig. 18-80 illustrates both the unsteered and steered behavior of an array of nine elements where the array length is twice the wavelength at the operating frequency. The angular interval between the half pressure points of the principal lobe in this example is about 30°.

The behavior of discrete element line arrays may be modified in a fashion similar to that employed with the continuous line array. Modifications are brought about by changing the amplitude and or the phase of the exciting signal applied to each element of the array. Digital signal processing allows the modifications themselves to be frequency dependent. When properly applied such processing can make an array segment's behavior nearly independent of the operating frequency. As an example, when amplitude tapering is applied to the nine-element array of Fig. 18-80 the improvements evidenced in Fig. 18-81 result.

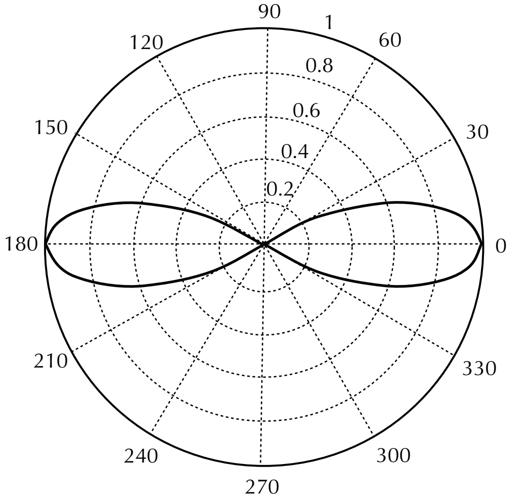


Unsteered

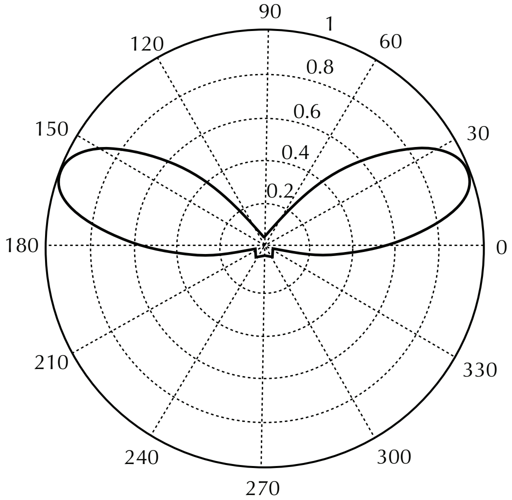


Steered +20°

Figure 18-80. Steering behavior of an array of nine elements.



Unsteered



Steered +20°

Figure 18-81. Steered nine-element array after amplitude tapering.

Amplitude tapering applied to the nine-element array greatly suppresses the side lobes formerly present and increases the beam width from 30° to 45°. As beneficial as amplitude tapering can be it must also be remarked that amplitude tapering also reduces the attainable pressure amplitude produced by the array. Linear amplitude tapering for example can bring about as much as a 6 dB reduction in attainable pressure amplitude as compared with a non-tapered array.

Arrays may be combined with the objective of producing coverage patterns with adjustable beam widths. Discrete line arrays of small transducers may be placed side by side and steered independently to produce results similar to those appearing in Fig. 18-82.

The pattern of Fig. 18-82 is obtained from arrays of nine elements placed side by side as closely as possible. The center array is unsteered, the left array is steered upward through a chosen angle, and the right array is steered downward through the same angle. The resulting beam width of the combination in this instance is 100°.

In summary, line arrays of discrete elements for employment in the far field operate best when separate lines are arranged for the low, mid, and high frequency ranges. Regardless of the frequency range, the length of the line determines the low frequency limit while the high frequency limit is determined by the minimum obtainable spacing between line elements. A minimum line length equal to λ is required to obtain directional control.

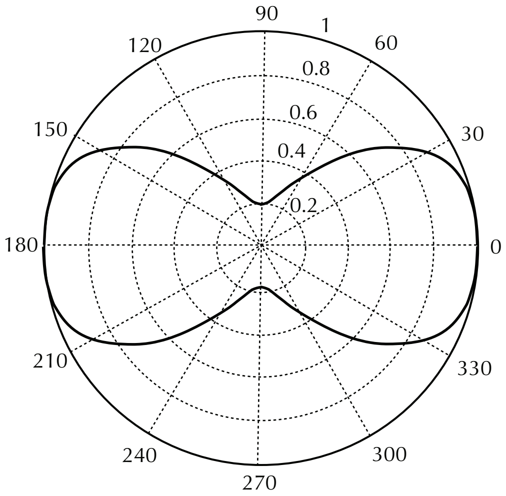


Figure 18-82. Array beam width shaping.

Constant directivity operation is possible by altering the line length inversely with the operating frequency by means of digital signal processing. Additional processing principally involving signal delay and amplitude adjustment versus line element position allows for beam steering as well as beam shape adjustment.

Line Arrays in the Near Field

Large line arrays of discrete elements designed for near field listening are usually composed of several full range loudspeakers, each of which is constructed similar to that displayed in Fig. 18-83.

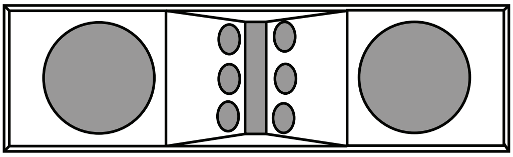


Figure 18-83. Full range loudspeaker module.

The typical module consists of two 15 in woofers housed in separate vented enclosures on the extreme left and right of the assembly. These enclosures border a mid-range horn featuring several co-entrantly mounted mid-range cone drivers. The mid-range horn has a co-axially mounted high frequency element. The high frequency element itself is designed to be an approximate continuous radiator with a height only slightly less than that of the total enclosure. The structure of the high frequency element approximates a cylindrical radiator and can occur in several possible forms. Its structure might be that of a planar diaphragm driving a vertical slot radiator, a wave guide driven by multiple compression drivers, or a continuous ribbon type tweeter.

When several of these modules are stacked one above the other the overall vertical extent may well be several meters in extent. The principal axis of a vertical array of these modules is the axis perpendicular to the line of the array passing through its center. The acoustic pressure along this axis falls off approximately at the rate of 3 dB per doubling of the distance for observation points in the near field along this axis. The extent of the near field is proportional to the frequency as indicated in the approximate relationship expressed in Eq. 18-137.

Consider an array that has a length of 7 m. An application of Eq. 18-137 predicts that the near field for an operating frequency of 5 kHz extends to over 350 m. The corresponding values at 500 Hz and 50 Hz would be 35 m and 3.5 m, respectively. Such an array when elevated above the stage would be capable of projecting intense mid and high frequency sound toward the distant nosebleed seats in a large arena or theatre. This statement is supported by the following simple analysis. The distant seating area is tiered. In a typical space the nearest seating may be at about 60 m with the most distant at about 80 m. This entire interval is in the near field zone for high frequencies where the distance attenuation rate is uniform at 3 dB per doubling of distance.

This is not the case for the mid frequencies. This interval is entirely in the far field for the mid-range. The median distance to the seating area is 70 m. In the interval from 35 m to 70 m, the mid-range is attenuated approximately 6 dB while the high range is attenuated approximately 3 dB. This can be readily compensated by a simple pre-emphasis of 3 dB applied to the mid-range.

The situation is not so happy for the low range. Between 3.5 m and 35 m, a decade, the low range is attenuated approximately 20 dB. In this same interval the mids and highs would have been attenuated by approximately half of this amount or only 10 dB. It is not likely that this large discrepancy can be adequately compensated without encountering power handling and headroom problems. The sound in these less expensive nosebleed seats can be quite intelligible while suffering from a lack of normal spectral balance.

Many practitioners employ a modification of the line array discussed above to provide coverage for the entire seating area of an arena. In this instance there is a straight line array segment that is usually tilted so that its principal axis is normal to the tiered seating plane of the distant nosebleed seats.

Attached to the bottom of this straight section is a curved section that is directed toward the floor seating area. This lower curved section is composed of modules identical to those of the straight section.

The curvature of the lower section constitutes a splay of the modules and affords a wide angular coverage as a result. The overall structure is termed a J array because its profile is similar to that of the letter for whom it is named. This lower portion of the array features individual drive and equalization to each of the modules constituting the curved section so as to produce a fairly uniform sound field over the entire floor seating space. The space attenuation rate for this curved section falls somewhere between that of a point source and a line source.

Distributed Systems

Distributed systems are employed in difficult acoustical spaces where speech intelligibility cannot be obtained by any of the foregoing techniques. Distributed systems raise the $\%AL_{CONS}$ by positioning sources of direct sound closer to the listeners. As a consequence the operating levels of the sources can be reduced thereby reducing the total acoustic power injected in the listening space. The net effect is an increase in the ratio of the direct sound level to the reverberate sound level, that in turn increases intelligibility.

Distributed systems can take many forms depending upon the conditions of the particular space involved. In spaces where the seating area is long with a small width, an inline system of two or more identical loudspeaker systems displaced along the central axis and operating with signal delay is a first consideration. Such a system is sketched in Fig. 18-84. Where it can be employed, this arrangement offers the advantage of maintaining source identification.

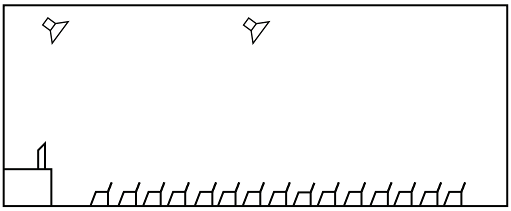


Figure 18-84. Inline array elevation view.

A related treatment is that which is applicable to fan shaped seating and is often encountered in some church construction. In this situation an often viable solution consists of a central source system attended by satellite systems displaced along successive radial arcs. Source identification is still maintained when appropriate signal delays are applied to the satellite loudspeakers. A typical plan view is sketched in Fig. 18-85. For equal area coverage by each satellite, the number of units employed must grow as the square of the radial distance from the origin of the fan.

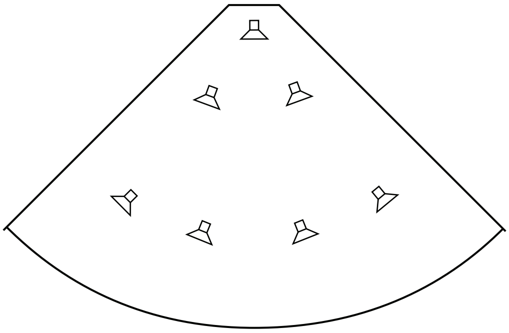


Figure 18-85. Satellite distributed system plan view.

The most often encountered distributed system is that of a pattern of overhead loudspeakers. The important considerations for such systems are pattern geometry, individual loudspeaker coverage angle, vertical distance between loudspeaker position and listening plane, and loudspeaker coverage pattern overlap. The starting point in the design of an overhead distributed system is the footprint that an individual loudspeaker has in the listening plane. Most loudspeakers employed in overhead systems have coverage patterns exhibiting cylindrical symmetry, i.e., the footprint in the listening plane is a circle whose radius can be calculated with the aid of Fig. 18-86.

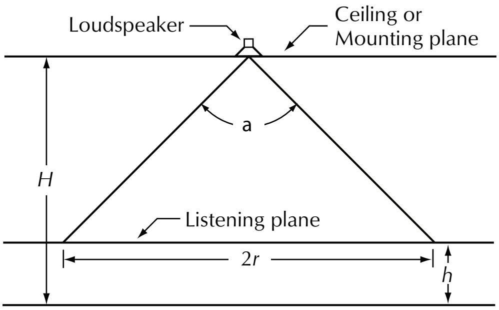


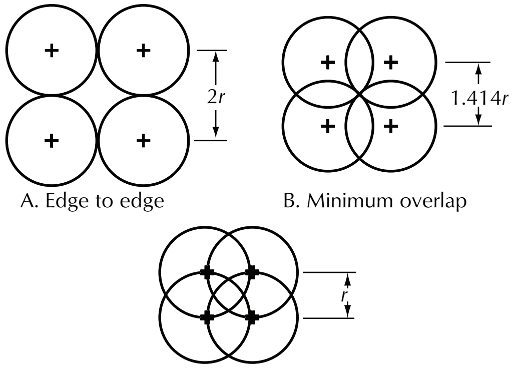
Figure 18-86. Single loudspeaker in an overhead distributed system.

The loudspeaker sizes employed in overhead systems are typically 5 inch, 8 inch, 8 inch co-axial, and 12 inch co-axial units. The coverage angle, α, in the critical speech range for the 5 inch as well as both of the co-axial units is usually taken as 90°

while that of the 8 inch non-co-axial is closer to 60°. The radius of the circular footprint of a single loudspeaker in the listening plane is calculated from

$$r = (H - h)\tan\left(\frac{\alpha}{2}\right) \qquad (18\text{-}151)$$

Loudspeaker layout patterns are based upon the geometrical properties of two regular polygons. Topological properties dictate that the chosen regular polygons be either a square or a hexagon. The implementation involving the square is simpler while that involving the hexagon affords a larger loudspeaker density attended by more uniform coverage. In both instances a choice must be made as to the degree of loudspeaker pattern overlap. The choice is usually made among the three presented in both Figs. 18-87 and 18-88 dealing with the square and hexagon, respectively.

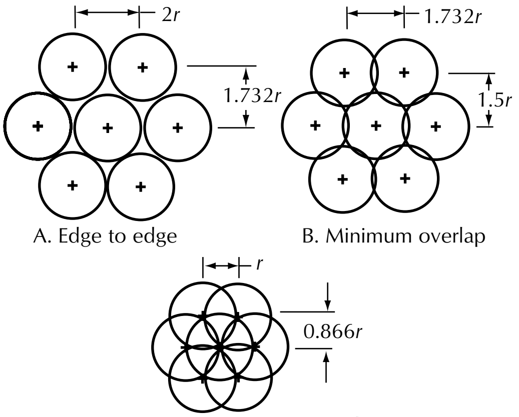


A. Edge to edge

B. Minimum overlap

C. Center to center

Figure 18-87. Overlap choices for a pattern based on a square cell.



A. Edge to edge

B. Minimum overlap

C. Center to center overlap

Figure 18-88. Overlap choices based on a hexagonal cell.

The dimensions of the basic cell employed depend upon the radius of the individual loudspeaker coverage pattern and the choice of overlap. The edge to edge choice is most economical with regard to the number of loudspeakers ultimately required while leaving gaps in the overall coverage. Quality installations employ either minimum overlap or center to center overlap. Of the two basic cells, the hexagonal structure provides higher loudspeaker density as illustrated in the following example.

A large meeting room has a width of 75 ft and a length of 125 ft with the loudspeaker mounting plane located 25 ft above ear level. The loudspeakers to be employed have a nominal coverage angle, α, of 90°. An application of Eq. 18-151 indicates that $r = 25$ ft. Taking uniformity of coverage as the chief goal, it is decided at the outset to employ the center to center overlap choice.

This being the case, the dimension of a square cell is 25 ft on edge. The dimensions of a hexagonal cell would be 25 ft and 21.65 ft. Layout patterns are to be generated for both basic cells. A possible layout procedure in which symmetry is assured consists of placing the chosen basic cell at the center of the space and attaching and/or overlaying additional cells up to the point where the overall pattern remains within the confines of the seating space. Fig. 18-89 presents two possible solutions obtained when square cells are employed.

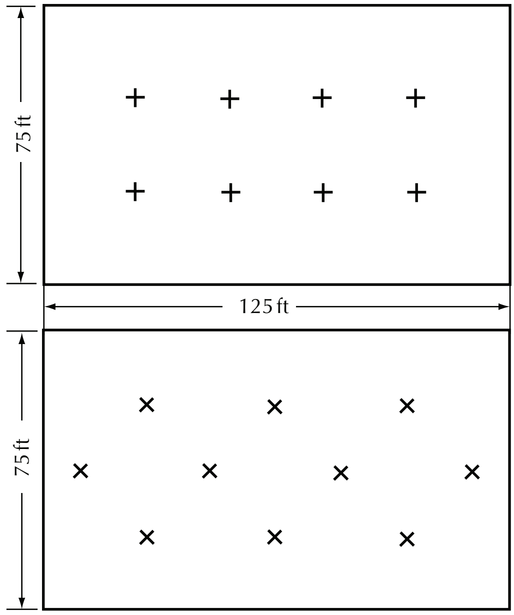


Figure 18-89. Loudspeaker positions employing square cells.

In the first solution the square cell, in its normal orientation, is placed in the center of the seating space and is replicated two times for a total of three cells. Further replications would place speakers on or beyond the boundaries. This solution involves a total of eight loudspeakers. The second solution presented in Fig. 18-89 involves a simple re-orientation of the basic square cell. In this instance the cell is rotated into the diamond or diagonal position.

Starting again with a rotated cell in the center only two replications are possible, yielding a total of three cells as before. In this instance, however, a total of ten loudspeakers fill the space thus improving the uniformity of coverage. Fig. 18-90 presents the two solutions obtained when the basic cell is a hexagon.

The first solution presented in Fig. 18-90 starts with a hexagonal cell of normal orientation placed at the center. Cells are replicated to both the right and left yielding a total of three cells. This solution gives a total of thirteen loudspeaker positions. This is a significant improvement in coverage as compared with either of the solutions based on a square cell. In the second solution presented in the Fig. 18-90 the center cell is in the rotated or diagonal position. Replication of this pattern again yields a total of 13 loudspeakers. This second solution yields slightly improved uniformity of coverage near the boundaries. Other procedures are possible for laying out the basic loudspeaker position patterns. Enerson has formulated algorithms that are useful aids in pattern organization.

Hybrid Arrays

Many theaters and churches have deep balcony overhangs beneath which there is no line of sight communication with a central cluster position. In such instances the central cluster is designed to cover the non-obscured seating spaces and a separate distributed system is designed to cover the obscured areas beneath the balcony. The distributed system is operated with a signal delay consistent with direct arrival from the central cluster. If the balcony is very deep it is necessary to zone the distributed array and to employ stepped signal delays appropriate to the zone positions.

Split Identical Sources

Unfortunately there exist many school and other small auditorium installations where identical loudspeakers are positioned symmetrically to the left and right of the stage. Such installations guarantee that all listeners other than those seated exactly on the centerline are subjected to a spectrum filled with comb filters. This type of installation is to be avoided where possible.

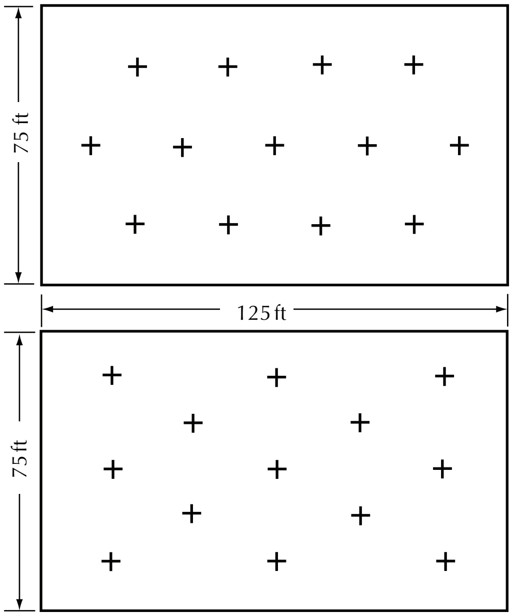


Figure 18-90. Loudspeaker positions employing hexagonal cells.

Comb filters always occur with spatially separated sources handling identical program material. Comb filters are produced by overhead distributed systems also. In this instance, though, there are several sources at different locations and the comb filters are very dense. A system consisting of N spatially separated sources has $[N(N-1)]/2$ distinct pairs.

For an overhead system of 10 sources there will be 45 distinct pairs of sources. The peaks formed from one pair of speakers often fill the notches caused by another pair of loudspeakers. The resulting spectral distortion is thus not as severe as that caused by a single pair of sources. Fig. 18-91 depicts the amplitude response through the critical speech intelligibility range in the direct sound field at a typical seating position not on the centerline in an auditorium having split sources.

As can be seen from Fig. 18-91 the notches are quite deep and the peaks are nearly 6 dB relative to those produced by a single source. Attenuating one source by 3 dB can produce a slight improvement. This produces the result displayed in Fig. 18-92.

The notches now are considerably less deep while the peaks are less than 5 dB. The price for this improvement is that one side of the auditorium will be less loud than the other side. All of this can be avoided through the employment of a single source suitably placed.

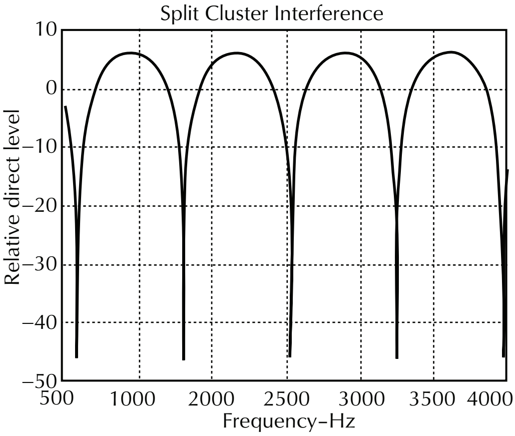


Figure 18-91. Comb filter response at a seat not on centerline.

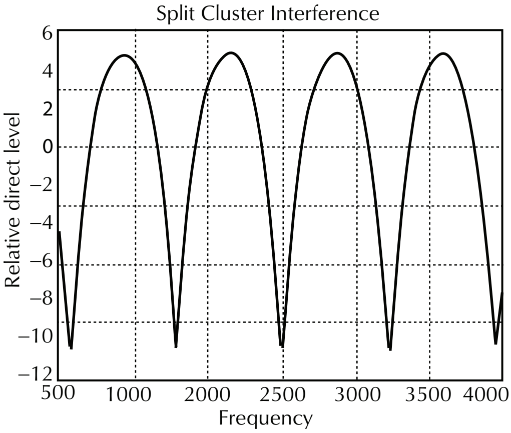


Figure 18-92. Result with one source attenuated 3 dB.

18.16 Vented Enclosure Bass Loudspeakers

Three types of vented enclosure bass loudspeakers will be covered here. The first type is the classic bass reflex loudspeaker to be followed by a closely related symmetrical bandpass subwoofer. The section will close with an analysis of an asymmetric bandpass subwoofer that has gained popularity in home theater systems.

A vented or ported loudspeaker is one in which the main loudspeaker enclosure is not sealed but rather has an opening or openings into the enclosure. A simple such classic structure is shown in Fig. 18-93. This particular structure is also called a bass reflex enclosure.

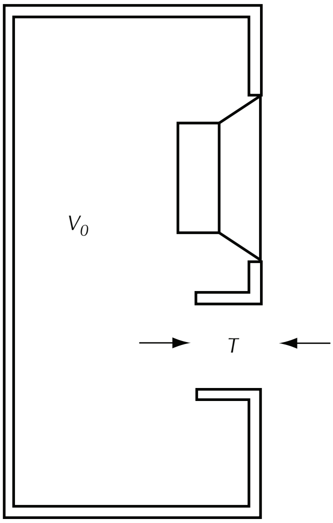


Figure 18-93. Cross section of a simple vented loudspeaker.

The volume available for free air in the enclosure, V_0, is the total interior volume of the enclosure less the volume occupied by the loudspeaker driver and less the volume occupied by the vent structure. The vent structure has an overall depth of T. The vent is preferably circular with a radius a_v, and an area S_v. In all of the following, the static air density will be represented by ρ_0 and the speed of sound by c. The properties of the vent can be listed as:

a. $M_v = \rho_0 S_v T$ = vent air mass.

b. $M_v' = M_v + \dfrac{\rho_0 8 S_v a_v}{3\pi}$ = vent air mass plus contribution from vent radiation reactance.

c. $K_v' = \dfrac{\rho_0 c^2 S_v^2}{V_0}$ = stiffness of vent.

d. $\omega_v = \sqrt{\dfrac{K_v'}{M_v'}}$ = vent resonant angular frequency.

e. $R_{rv} = \dfrac{\rho_0 S_v^2 \omega^2}{2\pi c}$ = vent radiation resistance.

f. $Z_{mv} = R_{rv} + j\left(\omega M_v' - \dfrac{K_v'}{\omega}\right)$ = vent mechanical impedance.

Similarly, the properties of the driver may be listed as:

a. $Z_e = R_e + j\omega L_e$ = blocked electrical impedance of driver.

b. a_d = effective radius of driver.

c. S_d = effective area of driver.

d. K_d = stiffness of driver in free air.

e. $K_d' = K_d + \dfrac{\rho_0 c^2 S_d^2}{V_0}$ = stiffness of driver in sealed enclosure of volume V_0.

f. M_d = effective moving mass of driver.

g. $M_d' = M_d + \dfrac{\rho_0 8 S_d a_d}{3\pi}$ = moving mass of driver including effect of radiation reactance.

h. R_{md} = mechanical resistance of driver suspension and surround.

i. $R_{rd} = \dfrac{\rho_0 S_d^2 \omega^2}{2\pi c}$ = radiation resistance of driver.

j. Bl = motor strength of driver.

k. $Q_t = \dfrac{\sqrt{K_d M_d'}}{R_{md} + \dfrac{(Bl)^2}{R_e}}$ = total quality factor of driver in free air.

l. $\omega_r = \sqrt{\dfrac{K_d}{M_d'}}$ = free air resonant angular frequency of driver.

m. $\omega_d = \sqrt{\dfrac{K_d'}{M_d'}}$ = resonant angular frequency of driver in sealed enclosure of volume

$V_0 . Z_{md} = R_{md} + R_{rd} + j\left(\omega M_d' - \dfrac{K_d'}{\omega}\right)$ = mechanical impedance of driver.

n. i = phasor describing voice coil current.

o. E = phasor description of drive voltage from source having negligible impedance.

The motions of the driver diaphragm and the air mass contained within the volume of the vent interact with each other. As an example, suppose for the moment that the air in the vent remains at rest. Further suppose that the driver diaphragm is displaced inward thus reducing the volume available for the air in the box. This would elevate the pressure in the box thus producing an outwardly directed force to act on the air in the vent. Similarly, one could suppose the driver to be at rest with the plug of air in the vent being displaced inward. This would cause an outwardly directed force to act on the driver diaphragm. The motions of the driver diaphragm and the air plug in the vent are coupled. The mutual coupling factor is given by $\mu = (\rho_0 c^2 S_d S_v)/V_0$ = vent-diaphragm displacement mutual coupling factor.

Having provided this framework it is now possible to write the equations of motion for both the driver and the plug of air contained in the vent.

$$Z_{md} u_d + \mu \frac{u_v}{j\omega} = Bli \qquad (18\text{-}152)$$

$$Z_{mv} u_v + \mu \frac{u_d}{j\omega} = 0 \qquad (18\text{-}153)$$

Eq. 18-153 can be solved immediately to obtain

$$u_v = -\frac{\mu}{Z_{mv}} \frac{u_d}{j\omega} \qquad (18\text{-}154)$$

This result can then be substituted into Eq. 18-152 to obtain

$$\left(Z_{md} + \frac{\mu^2}{Z_{mv}\omega^2}\right) u_d = Bli \qquad (18\text{-}155)$$

The voice coil current can now be written as

$$i = \frac{E - Blu_d}{Z_e} \qquad (18\text{-}156)$$

Combining Eqs. 18-155 and 18-156 leads to

$$u_d = \frac{\dfrac{BlE}{Z_e}}{Z_{md} + \dfrac{\mu^2}{Z_{mv}\omega^2} + \dfrac{(Bl)^2}{Z_e}} \qquad (18\text{-}157)$$

Combining Eqs. 18-154 and 18-157 produces

$$u_v = \left(\frac{-\mu}{Z_{mv}\, j\omega}\right) \frac{\dfrac{BlE}{Z_e}}{Z_{md} + \dfrac{\mu^2}{Z_{mv}\omega^2} + \dfrac{(Bl)^2}{Z_e}} \qquad (18\text{-}158)$$

The driver velocity and vent velocity Eqs. 18-157 and 18-158 are phasor quantities because the form taken for the driving emf is given by

$$E = E_m e^{j\omega t} \qquad (18\text{-}159)$$

where,

E_m = amplitude of driving emf in volts.

At low frequencies both the driver and the vent act as simple spherical sources radiating into a half space. The phasors representing the acoustic pressures produced on axis in the far field at a radial distance r are given by

$$p_d = \frac{\rho_0 j\omega S_d u_d}{2\pi r} \tag{18-160}$$

$$p_v = \frac{\rho_0 j\omega S_v u_v}{2\pi r} \tag{18-161}$$

Both of these quantities are complex. One determines the magnitudes of the individual pressures by taking the absolute magnitude of the corresponding complex quantity. The quantity of principal interest is the total pressure produced by the loudspeaker. This is obtained by first adding the two phasor expressions of Eqs. 18-160 and 18-161 to obtain the phasor expression for the total pressure

$$p_t = p_d + p_v \tag{18-162}$$

The amplitude of the total acoustic pressure is then found by taking the absolute magnitude of the complex quantity of Eq. 18-162

$$p_{mt} = |p_t| \tag{18-163}$$

The amplitude performance of the loudspeaker is obtained by plotting $20\,\mathrm{dB}\log(p_{mt}/0.00002\,\mathrm{Pa})$ versus $\log(\omega)$ over the frequency range of interest. The phase performance can be displayed by plotting the angle associated with the complex exponential statement of Eq. 18-162 versus $\log(\omega)$.

The foregoing equations are perfectly general and are applicable regardless of the choice of loudspeaker parameters, enclosure volume, and vent properties. As such, they furnish little or no guidance as to what choices need to be made for the various parameters in order to obtain a particular frequency response. If we neglect the radiation resistances and the voice coil inductance at the lowest frequencies it is possible to write Eq. 18-162 in the following form:

$$p_t = \frac{\dfrac{\rho_0 S_d E_m B l \omega^4}{2\pi r M_d' R_e \omega_0^4}}{\dfrac{\omega^4}{\omega_0^4} - \dfrac{j\omega^3}{Q_t\omega_0^3} - \dfrac{\omega_v^2 + \omega_d^2}{\omega_0^4}\omega^2 + \dfrac{j\omega_v^2\omega}{Q_t\omega_0^3} + 1} e^{j(\omega t - kr)} \tag{18-164}$$

where,

$$\omega_0 = \sqrt[4]{\frac{K_d K_v'}{M_v' M_d'}}.$$

The exponential term on the extreme right of Eq. 18-164 describes wave propagation, has a magnitude of one, and plays no role in determining the performance of the loudspeaker. Mathematically it describes the increasing phase lag as the radial distance increases. The shape of the frequency response is governed by the frequency factors in the numerator along with the coefficients associated with the polynomial in the denominator. At very low frequencies the denominator has a value of one while the numerator grows with the fourth power of the angular frequency. At higher frequencies the denominator magnitude is also growing with the fourth power of the frequency so the response is that of a fourth order high pass filter. Whether the response has steps, ripples, is very slow in reaching its final value, or is maximally flat is determined solely by the nature of the polynomial in the denominator. The polynomial structure that leads to the optimum transient response as well as the maximally flat amplitude response is that of the fourth order Butterworth. This is the response that will be detailed here. In order to obtain this response, the denominator must have the form

$$\frac{\omega^4}{\omega_0^4} - j\sqrt{4+2\sqrt{2}}\frac{\omega^3}{\omega_0^3} - [2+\sqrt{2}]\frac{\omega^2}{\omega_0^2} +$$
$$j[\sqrt{4+2\sqrt{2}}]\frac{\omega}{\omega_0} + 1$$

This stringent requirement will be satisfied if and only if $\omega_v = \omega_r = \omega_0$, $\omega_d = \omega_0\sqrt{1+\sqrt{2}}$, and

$$Q_t = \frac{1}{\sqrt{4+2\sqrt{2}}}$$

The physical significance of ω_0 is that this is the angular frequency for which the response is down by 3 dB. Note that the free air resonant angular frequency of the driver as well as the resonant angular frequency of the vent must be equal to ω_0. This means that the −3 dB point of the system occurs at the free air resonant frequency of the driver and that the vent must be designed to resonate at this frequency. Furthermore, the total quality factor of the driver in free air must be $Q_t = 1/\sqrt{4+2\sqrt{2}} = 0.383$. The interior volume of the enclosure is dictated by $\omega_d = \omega_0\sqrt{1+\sqrt{2}}$. This requirement leads to

$$V_0 = \frac{\rho_0 c^2 S_d^{\;2}}{K_d\sqrt{2}} = \frac{1.4 P_0 S_d^{\;2}}{K_d\sqrt{2}}$$

where,
P_0 is the static atmospheric pressure.

The requirement that the Butterworth response places on the driver total quality factor means that drivers should be specifically tailored for this particular application. Fortunately, the response is not extremely sensitive to variations in Q_t so reasonable tolerances in this value are allowed. In summary, to construct a vented enclosure with the Butterworth B4 alignment:

1. Select a driver whose free air resonance matches the system's desired −3 dB point and whose total quality factor is 0.38 ± 20%.
2. Determine the driver suspension stiffness and use this in determining required enclosure volume.
3. Pick a reasonable vent area and determine vent stiffness.
4. Determine vent air mass necessary to make vent resonate at the system's −3 dB point.
5. Use vent area and air mass to determine vent length *T*.

It may be necessary to iterate steps 3 through 5 to obtain reasonable vent dimensions.

6. Pick enclosure dimensions to provide required free air volume consistent with accommodating volume occupied by both driver and vent.
7. For large enclosures where standing waves may fall within the pass band of the system, employ dimension ratios that do not exacerbate standing waves.
8. The enclosure must be rigidly braced and the volume occupied by bracing cannot be allowed to detract from free air interior volume.

The above equations and procedures will now be employed to explore the performance of a system designed around the properties of a commercially available woofer. The given woofer parameters are:

a. $R_e = 6.6\,\Omega$.
b. $L_e = 0.003\,\text{H}$.
c. $f_r = 18.3\,\text{Hz}$.
d. $K_d = 1339\,\text{N/m}$.
e. $M_d' = 0.102\,\text{kg}$.
f. $Q_t = 0.379$.
g. $R_m = 4.07\,\text{kg/s}$.
h. $Bl = 13.28\,\text{Tm}$.
i. $S_d = 0.0847\,\text{m}^2$.

An application of the design equations yields the following derived values:

a. $V_0 = 0.537\,\text{m}^3$.
b. $K_v' = 118.7\,\text{N/m}$.
c. $M_v' = 0.00898\,\text{kg}$.
d. $S_v = 0.0212\,\text{m}^2$.
e. $a_v = 0.0821\,\text{m}$.
f. $T = 0.283\,\text{m}$.

The performance of the system is first explored by determining the amplitude response at one meter when the driving voltage has an amplitude of 1 volt. Curves are produced for the vent pressure response, driver pressure response, and the system pressure response, Fig.18-94.

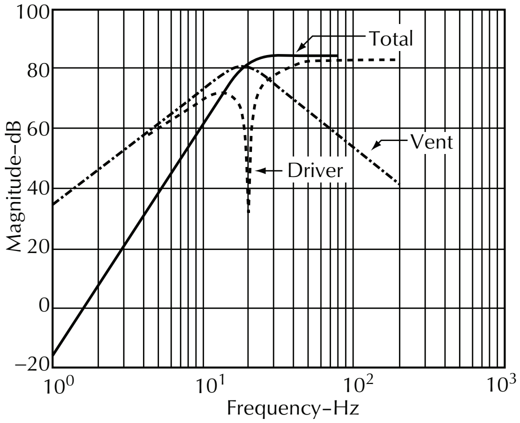


Figure 18-94. Amplitude performance of maximally flat vented bass loudspeaker system.

It is worthy of note that the vent response has a bandpass shape centered on the system's half power point located at 18.3 Hz and that the driver has a minimum output at this same point. Additionally, the driver output and the vent output interfere destructively for frequencies below the half power point. This is even more apparent upon examination of the separate phase responses presented in Fig. 18-95.

Note that for frequencies below the half power point the phase difference between the driver and the vent is uniformly π radians. They shift into phase as the frequency passes through the half power point and remain in phase for frequencies above the half power point.

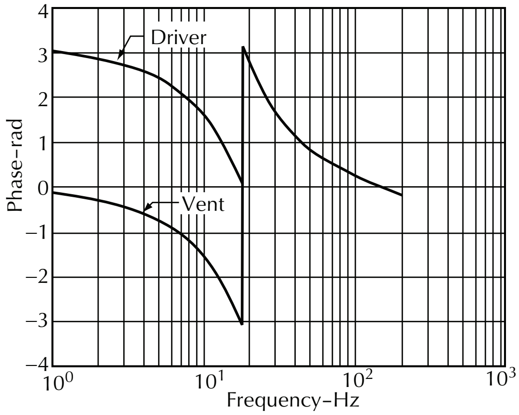


Figure 18-95. Phase response of vent and driver.

A final very important consideration is that of the displacement demanded of the driver. The displacement equation is obtained by dividing Eq. 18-157 by $j\omega$. One then takes the absolute magnitude of the result and plots this quantity as the frequency is allowed to vary over the range of interest. The result for the present maximally flat system appears in Fig. 18-96.

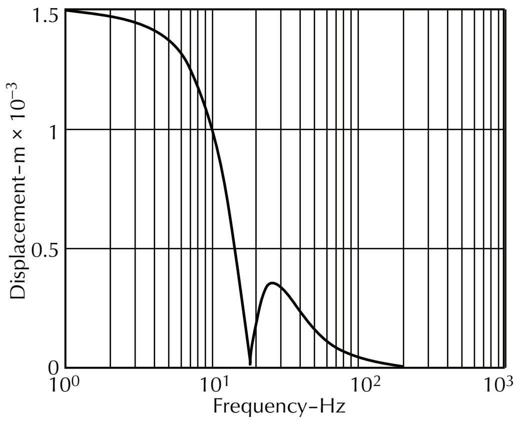


Figure 18-96. Driver displacement for one volt drive.

It should be noted that the maximum displacement occurs at the lowest drive frequency. The driver displacement at the half power point frequency is almost but not quite zero. At 1 Hz, the displacement amplitude is 1.5 × 10^{-3} m, 1.5 mm, or about 0.06 in. Curves of this type are important in determining the maximum allowable voltage drive as all drivers have linear displacement limits.

Equal Slope Bandpass Subwoofer

The design objective of bandpass subwoofers is the production of a device that radiates solely from the system vent or vents depending on the complexity of the design. The role of the driver in such devices is simply to excite air motion in the radiating vents. Fig. 18-97 depicts the simplest of such structures.

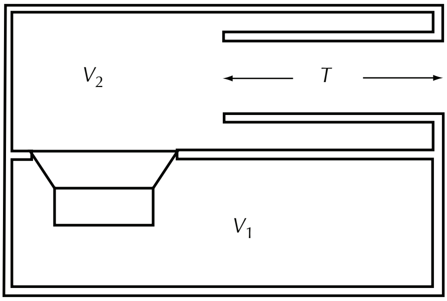


Figure 18-97. Simple bandpass subwoofer structure.

The properties of the vent can be listed as:

a. $M_v = \rho_0 S_v T$ = vent air mass.

b. $M_v' = M_v + \dfrac{\rho_0 8 S_v a_v}{3\pi}$ = vent air mass plus contribution from vent radiation reactance.

c. $K_v' = \dfrac{\rho_0 c^2 S_v^{\,2}}{V_2}$ = stiffness of vent.

d. $\omega_v = \sqrt{\dfrac{K_v'}{M_v'}}$ = vent resonant angular frequency.

e. $R_{rv} = \dfrac{\rho_0 S_v^{\,2} \omega^2}{2\pi c}$ = vent radiation resistance.

f. $Z_{mv} = R_{rv} + j\left(\omega M_v' - \dfrac{K_v'}{\omega}\right)$ = vent mechanical impedance.

Similarly, the properties of the driver may be listed as:

a. $Z_e = R_e + j\omega L_e$ = blocked electrical impedance of driver.

b. a_d = effective radius of driver.

c. S_d = effective area of driver.

d. K_d = stiffness of driver in free air.

e. $K_d' = K_d + \dfrac{\rho_0 c^2 S_d^{\,2}(V_1 + V_2)}{V_1 V_2}$ = stiffness of driver as mounted.

f. $K_d'' = K_d + \dfrac{\rho_0 c^2 S_d^{\,2}}{V_1}$ = stiffness of driver considering only rear volume.

g. M_d = effective moving mass of driver.

h. $\omega_d = \sqrt{\frac{K_d''}{M_d}}$ = resonant angular frequency of driver considering only rear volume.

i. R_{md} = mechanical resistance of driver suspension and surround.

j. Bl = motor strength of driver.

k. $Q_d = \frac{\sqrt{K_d''M_d}}{R_{md} + \frac{(Bl)^2}{R_e}}$ = quality factor of driver considering only rear volume.

l. $Z_{md} = R_{md} + j\left(\omega M_d - \frac{K_d'}{\omega}\right)$ = mechanical impedance of driver.

m. i = phasor describing voice coil current.

n. E = phasor description of drive voltage from source having negligible impedance.

o. $\mu = \frac{\rho_0 c^2 S_d S_v}{V_2}$ = vent-diaphragm displacement mutual coupling factor.

The equations of motion are:

$$Z_{md}u_d - \mu\frac{u_v}{j\omega} = Bli \tag{18-165}$$

$$Z_{mv}u_v - \mu\frac{u_d}{j\omega} = 0 \tag{18-166}$$

Note the reversal of sign of the mutual coupling term as compared with the bass reflex example. In this case, the vent is driven by the front of the driver rather than the rear. At very low frequencies, positive displacement of the driver produces a positive displacement of the vent unlike the former case.

As before the voice coil current is

$$i = \frac{E - Blu_d}{Z_e} \tag{18-167}$$

The procedure employed in the bass reflex case produces

$$u_d = \frac{\frac{BlE}{Z_e}}{Z_{md} + \frac{\mu^2}{Z_{mv}\omega^2} + \frac{(Bl)^2}{Z_e}} \tag{18-168}$$

$$u_v = \frac{\mu}{Z_{mv}j\omega}\frac{\frac{BlE}{Z_e}}{Z_{md} + \frac{\mu^2}{Z_{mv}\omega^2} + \frac{(Bl)^2}{Z_e}} \tag{18-169}$$

The phasor describing the acoustic pressure produced by the vent is obtained by solving Eq. 18-169 and then substituting the result into

$$p_v = \frac{\rho_0 j\omega S_v u_v}{2\pi r} \tag{18-170}$$

The pressure response of this system is that of a fourth order bandpass. In order for the denominator polynomial to conform to that of the fourth order Butterworth, the following conditions must be satisfied:

a. $\omega_0 = \omega_v = \omega_d$ = bandpass center angular frequency.

b. $Q_d = 0.383$.

c. $V_2 = \frac{\rho_0 c^2 S_d^2}{\sqrt{2}K_d''} = \frac{1.4P_0 S_d^2}{\sqrt{2}K_d''}$.

One must start with a driver whose free air resonance frequency is well below the desired pass band center frequency. The total quality factor of the driver in free air must also be well below 0.383. The volume V_1 is then determined from

$$V_1 = \frac{1.4P_0 S_d^2}{\omega_0^2 M_d - K_d}$$

Knowing the value of V_2, it is possible to determine the vent area and depth by following the procedure employed in the bass reflex example.

This procedure was applied to a small driver whose parameters met the above criterion with regard to free air resonance frequency but whose free air total quality factor was too high. This is often the case with small drivers so the performance might be more typical of actual practice than the theoretical ideal. The given parameters were:

a. $K_d = 1088\,\text{N/m}$.

b. $M_d = 0.0199\,\text{kg}$.

c. $a_d = 0.0847\,\text{m}$.

d. $R_e = 6.8\,\Omega$.

e. $L_e = 0.001\,\text{H}$.

f. $R_m = 3\,\text{kg/s}$.

g. $Bl = 7.94\,\text{Tm}$.

The calculated parameters are:

a. $V_1 = 0.0469\,\text{m}^3$.

b. $V_2 = 0.0194\,\text{m}^3$.

c. $T = 0.325\,\text{m}$.

The pressure amplitude response for a distance of one meter with a drive voltage amplitude of 1 volt is given in Fig. 18-98.

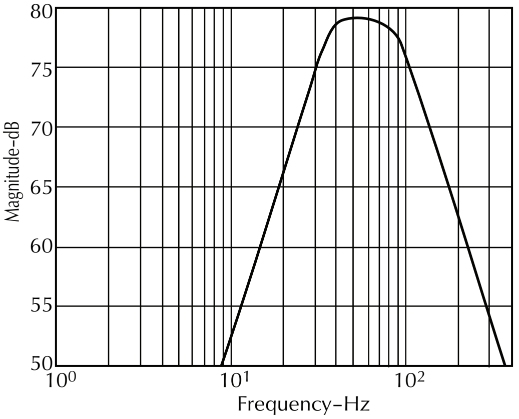


Figure 18-98. Amplitude response of single vent subwoofer.

The slopes are symmetrical on either side of band center and are 12 dB/octave. The bandpass extends from just over 30 Hz to just over 100 Hz. The response is not as smooth across the top as it should be because with this driver Q_d is 50% larger than it should be to generate a true maximally flat performance. The displacement required of the driver in producing the above vent response is equally important. This curve appears as Fig. 18-99.

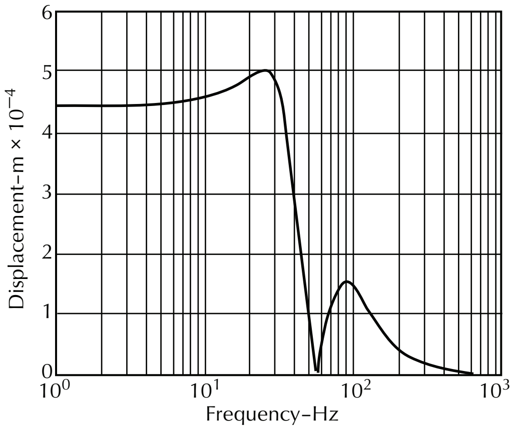


Figure 18-99. Driver displacement versus frequency with a voltage amplitude of one volt.

The peak displacement at any frequency with a one volt drive amplitude is 0.5 mm. This driver can tolerate a displacement of ten times this value. The system should be capable then of producing a pressure response curve elevated by 20 dB more than depicted in Fig. 18-98.

Unequal Slope Bandpass Subwoofer

A popular subwoofer for home theater systems features a small driver whose effective radius is a little over 5 cm while incorporating two vents. One vent is for the air volume in front of the driver as in the equal slope bandpass device discussed above. The second vent is for the air volume to the rear of the driver. The area of each of the radiating vents is equal to the effective area of the driver. The parameters for the system are:

a. R_{md} = mechanical resistance of driver = 1.33 kg/s.

b. M_d = effective moving mass of driver = 0.0133 kg.

c. K_d = stiffness of driver = 1261 N/m.

d. S_d = effective driver area = $0.00894\,\text{m}^2$.

e. V_1 = rear of driver air volume = $0.006\,\text{m}^3$.

f. V_2 = front of driver air volume = $0.0015\,\text{m}^3$.

g. Bl = driver motor strength = 6.6 Tm.

h. R_e = voice coil resistance = 3.1 Ω.

i. L_e = voice coil self-inductance = 0.00128 H.

j. S_{v1} = rear vent area = $0.00894\,\text{m}^2$.

k. S_{v2} = front vent area = $0.00894\,\text{m}^2$.

l. M_{v1}' = effective moving mass of rear vent = 0.0234 kg.

m. M_{v2}' = effective moving mass of front vent = 0.0143 kg.

The electroacoustic properties are:

a. $Z_e = R_e + j\omega L_e$.

b. $R_{r1} = \dfrac{\rho_0 S_{v1}{}^2 \omega^2}{2\pi c}$ = rear vent radiation resistance.

c. $R_{r2} = \dfrac{\rho_0 S_{v2}{}^2 \omega^2}{2\pi c}$ = front vent radiation resistance.

d. $$Z_{md} = \left[R_{md} + j\left(\omega M_d - \frac{K_d}{\omega}\right) + \frac{\rho_0 c^2 S_d{}^2 (V_1 + V_2)}{V_1 V_2 j\omega} + \frac{(Bl)^2}{Z_e}\right] =$$ driver mechanical impedance.

e. $\mu_1 = \frac{\rho_0 c^2 S_{v1} S_d}{V_1}$ = rear mutual coupling factor.

f. $\mu_2 = \frac{\rho_0 c^2 S_{v2} S_d}{V_2}$ = front mutual coupling factor.

g. $Z_{m1} = R_{r1} + j\left(\omega M_{v1}' - \frac{\rho_0 c^2 S_{v1}^2}{V_1 \omega}\right)$ = rear vent mechanical impedance.

h. $Z_{m2} = R_{r2} + j\left(\omega M_{v2}' - \frac{\rho_0 c^2 S_{v2}^2}{V_2 \omega}\right)$ = front vent mechanical impedance.

At this point it is possible to write the equations of motion.

$$Z_{md} u_d - \frac{\mu_2}{j\omega} u_{v2} + \frac{\mu_1}{j\omega} u_{v1} = \frac{BlE}{Z_e} \tag{18-171}$$

$$Z_{m1} u_{v1} + \frac{\mu_1}{j\omega} u_d = 0 \tag{18-172}$$

$$Z_{m2} u_{v2} - \frac{\mu_2}{j\omega} u_d = 0 \tag{18-173}$$

where,
u_d = driver diaphragm velocity,
u_{v1} = rear vent particle velocity,
u_{v2} = front vent particle velocity.

Eqs. 18-171 through 18-173 constitute a system of linearly independent equations that can be solved simultaneously for the three unknown velocity variables. The results are:

$$u_d = \frac{BlE_m e^{j\omega t}}{Z_e} \times \frac{Z_{m1} Z_{m2}}{\Delta} \tag{18-174}$$

$$u_{v1} = \frac{-BlE_m e^{j\omega t}}{Z_e} \times \frac{\mu_1}{j\omega} \times \frac{Z_{m2}}{\Delta} \tag{18-175}$$

$$u_{v2} = \frac{BlE_m e^{j\omega t}}{Z_e} \times \frac{\mu_2}{j\omega} \times \frac{Z_{m1}}{\Delta} \tag{18-176}$$

where,

$$\Delta = Z_{md} Z_{m1} Z_{m2} + \frac{\mu_1^2}{\omega^2} Z_{m2} + \frac{\mu_2^2}{\omega^2} Z_{m1}.$$

The phasor expression for the acoustic pressure produced by the system is then

$$p_t = \frac{\rho_0 j\omega}{2\pi r}(S_{v1} u_{v1} + S_{v2} u_{v2}) \tag{18-177}$$

The amplitude response of the system is obtained by plotting $20\log(|p_t|/0.00002\ \text{Pa})$ versus $\log(\omega)$or $\log(f)$. Such a plot appears in Fig. 18-100.

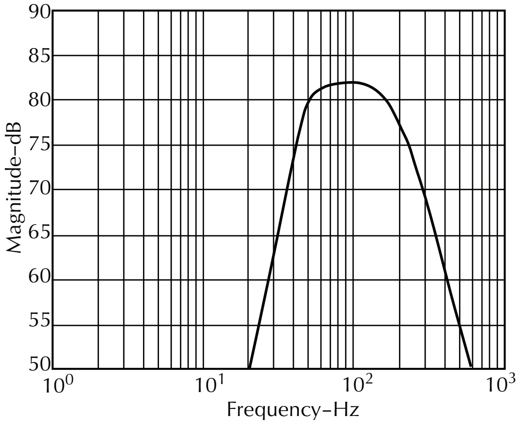


Figure 18-100. One volt one meter amplitude response of unequal slope subwoofer.

The driver displacement required for the pressure performance of Fig. 18-100 appears in Fig. 18-101.

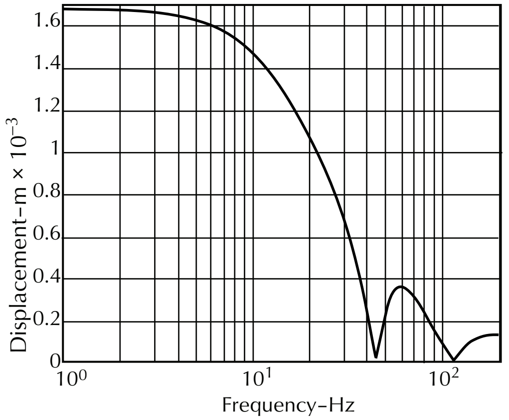


Figure 18-101. Driver displacement required by unequal slope subwoofer.

The half power points of the amplitude response curve occur at about 45 Hz and 180 Hz. The low frequency slope rate is 24 dB/octave while the high frequency slope rate is 18 dB/octave. This performance is exceptional considering the size of the driver involved. The displacement curve, however, requires a displacement maximum of over 1.6 mm and drivers of this size hardly exceed 3 mm max. The maximum output of the system is thus quite limited unless the drive signal is high passed above 10 Hz. Doing this, of course, would unfavorably impact the low frequency half power point.

18.17 Large Signal Behavior of Loudspeakers

The loudspeaker models presented earlier in the present chapter accurately describe a loudspeaker's performance under linear and time invariant conditions. These conditions exist when a loudspeaker is being driven by signals sufficiently small that the displacements experienced by the voice coil fall in a range that provoke only linear restoring forces from the loudspeaker suspension. Continuous operation with small signals will produce an elevated voice coil temperature through the heating effect of the small currents involved. This in turn leads to a slightly higher voice coil resistance that remains relatively stable as long as operation with small signals persists. This is accounted for in small signal parameter measurement procedure by exercising the loudspeaker before parameter measurement takes place. In this small signal regime the motor strength Bl and the voice coil self inductance L_e can be treated as constant parameters and as the voice coil currents are small any magnetic hysteresis effects in the magnetic structure are negligible. None of the above statements remain true when the loudspeaker operation involves operation with large signals where now the stiffness of the suspension, the motor strength, and the voice coil self inductance all become functions of the amount of voice coil displacement. Under such conditions the loudspeaker is operating nonlinearly. Nonlinearity of operation is evidenced by a distorted acoustical output wherein frequency components appear in the output that were not present in the exciting signal at the input to the loudspeaker. Furthermore, the loudspeaker operation is no longer time invariant as the voice coil temperature as well as magnetic hysteresis effects through the elevated voice coil currents involved depend upon the previous time history of the exciting signals. Furthermore, the type of loudspeaker mounting or enclosure may play a much more significant role under large signal conditions. If a cone type loudspeaker is mounted in a sealed enclosure of small interior volume large cone displacements will provoke nonlinear behavior of the air trapped in the enclosure while for vented enclosures turbulence may occur in the vent air flow.

Practically all workers in audio at one time or another have performed distortion measurements associated with transducers while a smaller number have done significant work toward analysis and reduction of distortion properties. In the last decade or so one worker merits particular mention. Wolfgang Klippel of Dresden, Germany during this period has worked almost exclusively on the development of a system of analyzers for measuring and identifying the sources of distortion in loudspeakers and other transducers while operating under large signal conditions. Many of Klippel's papers as well as references to the work of others are available for download from Klippel's company website www.klippel.de.

Bibliography

A. H. Benade and E. V. Jansson. "On Plane and Spherical Waves in Horns with Nonuniform Flare. 1: Theory of Radiation, Resonance Frequencies, and Mode Conversion," *Acustica*, Vol. 31, No. 2. (1974) pp. 80-98.

Leo L. Beranek. *Acoustics*. New York: McGraw-Hill, 1954.

C. Enerson. "Distributed System Pattern Analysis," *Syn-Aud-Con Tech Topics*, Vol. 5, No. 1 (1977).

Earl Geddes. "Acoustic Waveguide Theory," *J. Audio Eng. Soc.*, Vol. 37, No. 7 (1989).

D. W. Gunness and N. D. Butler. "Implementation of a Wide-Bandwidth, Digitally Steered Array." Paper presented at the 115th Convention of the Audio Engineering Society (2003).

David W. Gunness and Ryan J. Mihelich. "Loudspeaker Acoustic Field Calculations with Application to Directional Response Measurement." Paper presented at the 109th Convention of the Audio Engineering Society, Los Angeles (September 22-25, 2000).

E. V. Jansson and A. H. Benade. "On Plane and Spherical Waves in Horns with Nonuniform Flare. 2: Prediction and Measurements of Resonance Frequencies and Radiation Losses," *Acustica*, Vol. 31, No. 4. (1974) pp. 185-202.

D. B. Keele, Jr. "Low-Frequency Loudspeaker Assessment by Nearfield Sound-Pressure Measurement," *J. Audio Eng. Soc.*, Vol. 22 (May 1974), pp. 330-338.

Lawrence E. Kinsler, Austin Frey, et. al. *Fundamentals of Acoustics*, 4th ed. New York: John Wiley and Sons, 2000.

Helmut Krüger. “Mechanical Equalising—On the Design of Loudspeaker Diaphragms.” Paper presented at the 86th Convention of the Audio Engineering Society, Hamburg (March 7-10, 1989), Preprint No. 2775.

P. M. Morse. *Vibration and Sound*, 2nd ed. New York: McGraw-Hill, 1948.

Allan D. Pierce. *Acoustics: An Introduction to Its Physical Principles and Applications,* New York: Mc-Graw-Hill, 1981.

Chester W. Rice and Edward W. Kellogg, “Notes on the Development of a New Type of Hornless Loudspeaker.” Paper presented at the Spring Convention of A.I.E.E., St. Louis (April 13-17, 1925).

Chapter 19

Power Ratings for Amplifiers and Loudspeakers

by Pat Brown

19.1 Loudspeaker Power Ratings 425
- The Amplifier/Loudspeaker Interface 425
- The Power Rating 428

19.2 Active Loudspeaker Systems 428

19.3 Non-Linear Operation 428
- The Voltage Rating 429
- Mechanical Limit Testing 429
- Conclusion 429

19.4 The Amplifier as a Voltage Source 429

19.5 The Equivalent Amplifier Size–*EAS* 430

19.6 Power from a Voltage Source 431

19.7 Burst Testing 433

19.8 Power Rating Possibilities 434
- Real World Power Generation 434
- Things Amplifiers Hate 435
- Burst Testing Results 435
- Conclusion 435

19.9 Putting It All Together 435
- A Warning 435
- A Common Misconception 436
- The Input Sensitivity Control 436
- The 30/30 Guideline 437

19.10 Multi-way Loudspeakers 438
- Fixed Gain Settings 439
- Powered Loudspeakers 439

19.11 System Gain Structure 439

19.12 Combining *MIV* and *EAS* 439

Power ratings enjoy two distinctions in the field of sound system engineering.

1. From a technical perspective, a proper power rating yields the deepest insight into the performance of the "black boxes" that make up our sound systems. It establishes the relationship between voltage, current and impedance with regard to the input and output signal to a device, allowing us to account for all of input and output energy. From that perspective, we should have a power rating, expressed as a level in either dBm or dBW, for every device in the signal chain. But,
2. Power ratings are among the most abused and misunderstood specifications in audio. They are almost never measured correctly, and they are seldom presented in a way that can be validated.

The power ratings of amplifiers and loudspeakers in today's audio marketplace are usually so ambiguous as to be useless to the sound system designer. I am not a fan of power ratings. Claims such as "the amplifier is putting out 1000 watts" or "the loudspeaker needs 500 watts" are almost never correct, or even meaningful. Such loose usage is the norm rather than the exception, and has diminished the usefulness amplifier and loudspeaker power ratings to the audio practitioner. It is my intent for this chapter to connect the dots for those who wish to understand the theory behind the numbers, and to equip audio practitioners with what they need to design and calibrate sound systems, namely what will be read on a true-rms voltmeter at component interfaces. This information must either be measured, or mined from the power rating of the product. In order to understand how to extract this voltage from a power rating, we must look at its role in establishing the power rating. Any power rating that cannot yield the correct voltage at an interface is useless.

I will show some shortcuts to finding these interface voltages that can allow the use of power ratings to be avoided altogether in some cases. This does not detract from the technical validity of the power rating, but it can allow us to reserve it for in-depth investigations rather than field work.

My objective is not to replace any presently existing standards, but rather to provide safe guidance through a subject that is often very confusing to beginners and veterans alike in sound system design and installation. The suggested procedures are appealing for several reasons. First, and perhaps foremost, only the most familiar basic instrumentation and measurements are required in the setup procedures. Secondly, safeguards are incorporated to protect against damage to both loudspeakers as well as amplifiers. Thirdly, indicators of system satisfactory performance or lack thereof are clearly identified.

19.1 Loudspeaker Power Ratings

The role of the amplifier is to provide the voltage and current needed by the loudspeaker to produce the desired L_P at the listener. Voltage times current equals power. Once the loudspeaker's voltage needs and limitations are determined, an appropriate amplifier can be selected. Since much of the applied power is converted to heat, the loudspeaker has a power rating which describes the maximum electrical power that it can continuously dissipate while it is producing acoustical power. Exceed this power limit, and the loudspeaker will be permanently damaged.

The power rating need only be considered when there is a chance that it will be exceeded. It is good system design practice to stay well below the power ratings of both amplifiers and loudspeakers. This assures optimal performance and longer life.

The following describes the method that I use to determine loudspeaker power ratings. I developed it for testing loudspeaker power ratings for manufacturers. Key to the method is the use of the real-time transfer function to assess when loudspeaker's response is changing due to the heat build-up. This allows a limit to be established without destroying the loudspeaker. An understanding of how the test is performed and the specification established will provide some insights into how to avoid thermal damage in the field, as well as how to select an appropriate amplifier to drive the loudspeaker.

19.1.1 The Amplifier/Loudspeaker Interface

Amplifiers are interfaced with loudspeakers using a constant voltage interface. The power flow to the loudspeaker will predictably track the applied voltage under normal, linear operating conditions. When this voltage is applied to the loudspeaker's terminals, current will flow as determined by the impedance of the loudspeaker. A loudspeaker power rating requires the determination of the maximum continuous applicable rms voltage to the loudspeaker that will not cause thermal damage. I will call this the Maximum Input Voltage–*MIV.* The power rating is calculated from the *MIV* and rated impedance using the power equation.

$$W = \frac{MIV^2}{R} \qquad (19\text{-}1)$$

where,

W is the continuous power rating in watts,

R is the rated impedance of the loudspeaker, simplified to an equivalent resistance,

MIV is the maximum rms voltage to the loudspeaker.

The Method

To find the MIV of the loudspeaker, the following parameters must be monitored.

1. The drive voltage to the loudspeaker.
2. The output voltage of the power amplifier.
3. The axial frequency response magnitude of the loudspeaker.

Various types of broad band stimuli can be used. Pink noise is a logical choice. Some standards shape the spectrum of the pink noise to more closely resemble musical waveforms. The IEC and EIA noise spectra are shown in Fig. 19-1.

The noise can also be clipped to allow a higher rms voltage to be produced by a given amplifier size. Most standards mandate a 2-to-1 voltage ratio (crest factor = 6 dB), which requires severe clipping of the noise signal. Pink noise has an approximate 4:1 peak-to-rms ratio (12 dB crest factor), which emulates slightly clipped (or peak-limited) program material. IEC and EIA noise have a 2:1 peak-to-rms ratio (6 dB crest factor). The spectral shaping and clipping, once performed by a passive network, can be performed in any wave editor. Of course, the clipped waveform is distressing to listen to, but this is a heat test, not a sound quality test. The original motivation for this "pre-clipping" of the test waveform was to allow a higher rms voltage to be achieved prior to clipping the amplifier. In this age of amplifiers with 200 V rails, the pre-clipping is generally considered to be unnecessary and future standards will not likely mandate it. It has little if any effect on the MIV, which is based on the rms voltage.

The relationship between the drive voltage to the loudspeaker and the axial frequency response magnitude can be monitored on a two-channel analyzer that has the ability to subtract the two responses and display their difference in real time. Fig. 19-2 shows the test setup.

Under normal operating conditions all three parameters have a linear relationship—a 1 dB drive voltage increase produces a 1 dB increase of output voltage from the amplifier, as well as a 1 dB increase in the axial L_P from the loudspeaker. The objective of the MIV test is to determine when this relationship becomes non-linear due to heating of the loudspeaker.

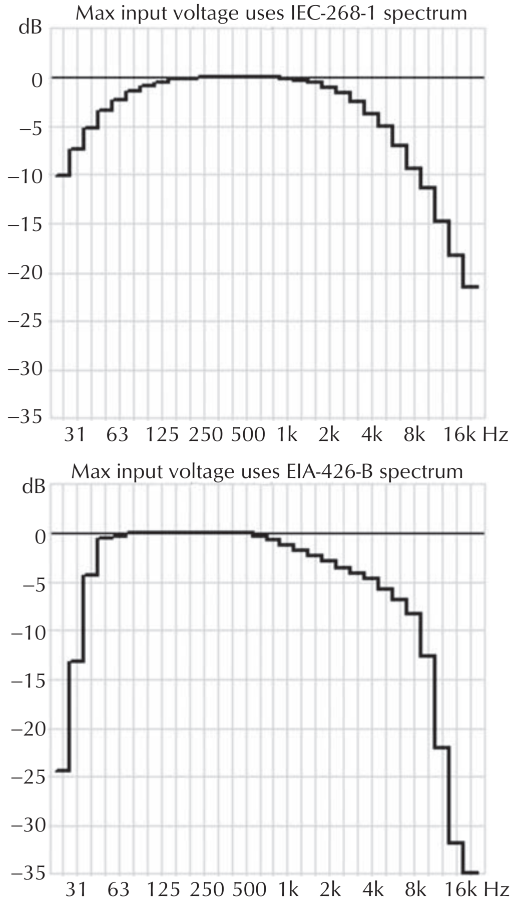


Figure 19-1. Test signal spectra for determining Maximum Input Voltage.

A Starting Point

I usually start by applying 3 V_{rms} to the loudspeaker. This is the approximate voltage used for the sensitivity test, and produces about 1 W into an 8 Ω resistive load. The analyzer displays the loudspeaker's response. This response is stored as a reference, and the analyzer is placed in "subtract" mode. This produces a straight line on the screen, indicating that there is no difference between the ongoing measurement and the response stored in memory. This line will remain flat so long as nothing causes a difference between the stored response and ongoing measurement.

The drive voltage is increased by turning up the generator. Since the analyzer sees both the reference signal and the loudspeaker response increase, the line remains flat because the same change

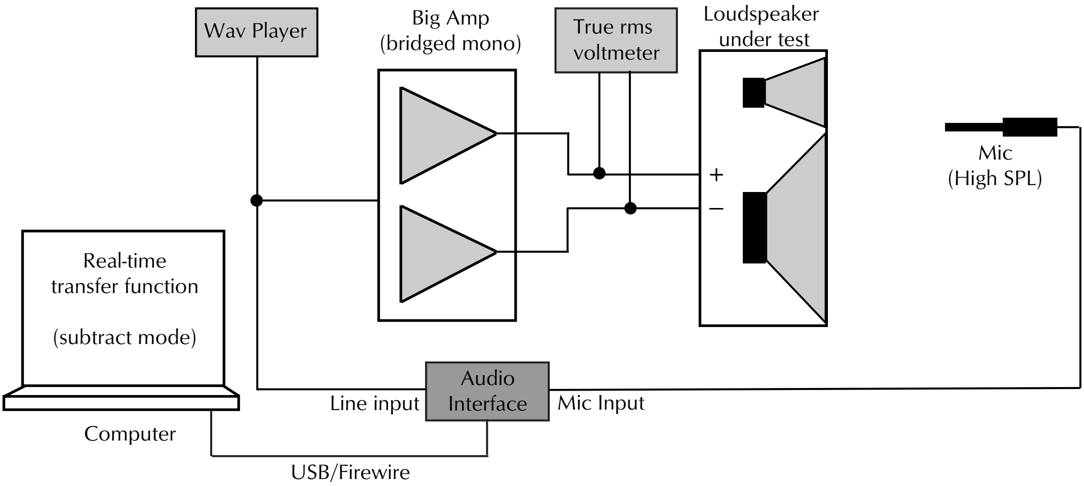


Figure 19-2. Setup for measuring the *MIV*.

happened to both signals. This demonstrates that the relationship between the two signals is linear, or 1:1.

Power Compression

The resistive portion of the loudspeaker's impedance dissipates electrical power in the form of heat. Loudspeakers can become very hot during use. Transducer designers use various mechanisms to keep them cool, including air flow and ferrous fluids.

As the heat increases the resistance of the voice coil wire increases, resulting in less current draw from the amplifier. This creates a non-linear relationship between the drive voltage and the resultant acoustic power. This change can be observed in the axial response of the loudspeaker. In short, power compression is occurring when some portion of the loudspeaker's frequency response magnitude no longer tracks the level increase of the applied voltage. Of course, it is important to make sure that some other part of the chain is not being overdriven, such as the measurement microphone.

For single transducers, the effect is usually broadband. For multi-way loudspeaker systems and for those with passive crossovers, it may be frequency-dependent. Which is unimportant. By monitoring the loudspeaker's response in the frequency domain, the frequency-dependence can be observed in real time.

Turn It Up!

The drive voltage to the amplifier is increased at regular time intervals. A 3 dB increase at one minute intervals is a good general approach. At the onset of power compression, the trace will start to drop in level, indicating that the output of the loudspeaker is no longer tracking the increase in drive voltage. A relationship that was once linear is now non-linear due to the heat produced in the loudspeaker's voice coil.

When To Stop?

There is no clear directive on when to conclude the test. Even the standards are ambiguous in this regard. The purist might say that the test is over at the onset of power compression, as indicated by a 1 dB change to any one-third octave band in the loudspeaker's response curve. I wouldn't disagree, except that the *MIV* and resultant power rating based on this criteria are quite conservative. Since the uninformed often purchase loudspeakers on the basis of their power rating alone, conservative ratings are shunned by manufacturers since they may produce a disadvantage in the marketplace. The influence of marketing forces on published ratings cannot be ignored.

One could ask "How hard can I drive the loudspeaker before its response is permanently changed?" After testing hundreds of units, I would suggest that the answer is about 3 dB to any one-third octave band, since the failure rate at 10 dB of change approaches 100% and the failure rate at 6 dB of change is still unacceptably high. Failures at a 3 dB change are rare, and the loudspeaker system typically recovers to its low power response when the drive voltage is reduced.

In a perfect world, we would like to know:

1. The voltage at the onset of power compression.

2. The voltage that produces 3 dB of power compression in any one-third octave band in the passband of the loudspeaker.
3. The voltage at which thermal failure occurs, which unfortunately requires a destructive test.

Of the three, item 2 is the most important. I conclude the test at a 3 dB response change to establish the *MIV*. The loudspeaker is driven for 3 minutes at this level (the length of a typical song). The drive voltage is then reduced to the original 3 V_{rms} drive. If no damage occurred, the analyzer trace should return to the flat line response after a few minutes.

If the loudspeaker's sensitivity is measured at 2.83 V_{rms} @ 1 m and the *MIV* is determined as described, the maximum L_P at one meter can be approximated by

$$dB = 20\log\left(\frac{MIV}{3}\right) + Sens_{AVG} \qquad (19\text{-}2)$$

This can be extrapolated to any listener distance using the inverse-square law.

19.1.2 The Power Rating

The *MIV* can be directly read from the true-rms voltmeter. How many watts is this? It depends on the loudspeaker's impedance. We will use the rated impedance of the loudspeaker along with the *MIV* to determine the power rating. Remember that this approach treats the loudspeaker like a resistor, simplifying the calculation.

We can simplify the complex impedance to an equivalent resistance, and even round to a typical value (i.e. 8 Ω). This is because once the *MIV* is known, the power rating is somewhat superfluous. We need to present the power rating in a form that allows the *MIV* to be determined by calculation, so whatever impedance is used basically falls out of the equation, so long as the same value is used in the establishment of the power rating as is used to calculate the *MIV* from it.

These extra steps are required due to the insistence of our industry on the use of power ratings, even for a constant voltage interface. This may change in the future.

19.2 Active Loudspeaker Systems

The same procedure can be used to test self-powered loudspeakers. The only difference is that the *MIV* will be a line level signal. The *MIV* is monitored at the input rather than at the output of the amplifier. This method can also be used on passive loudspeakers that are packaged as a system with amplifiers and signal processing, Fig. 19-3.

19.3 Non-Linear Operation

It is becoming increasingly common for a loudspeaker system response to become nonlinear by design as its thermal or mechanical limits are approached. This may be due to the use of electronic compression or limiting, either analog or digital, or by the use of a passive network within the loudspeaker enclosure. The test method previously

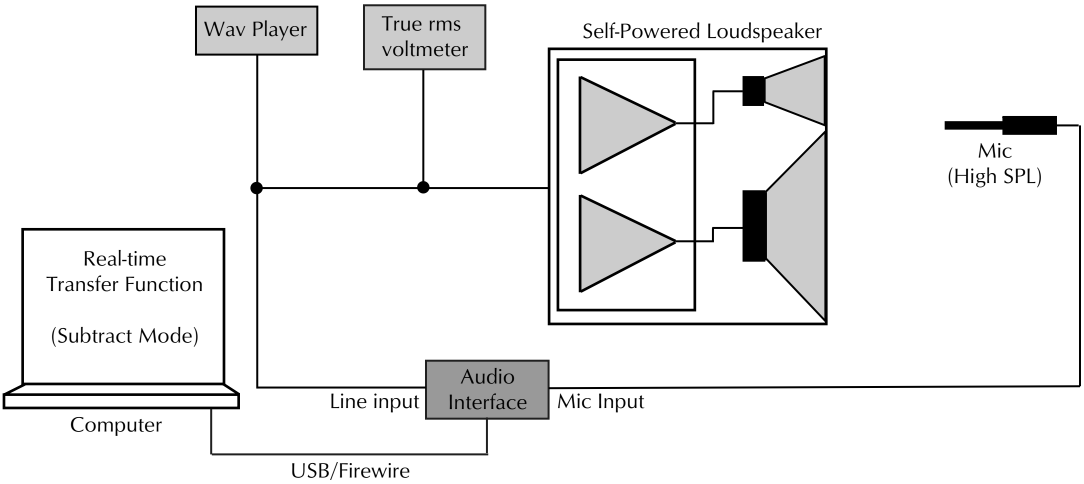


Figure 19-3. *MIV* test setup for internally powered loudspeaker.

described cannot distinguish between power compression due to heat build-up, and intentional compression designed to yield higher L_P from the system. The *MIV* can be tested as previously described to produce the maximum voltage for linear output. It can then be increased beyond this point to determine a new *MIV* using different failure criteria, such as a 1 dB change to the total L_P vs. the applied voltage. This will typically be several dB higher than *MIV* resulting from the one-third octave criteria. In effect, the signal processing is impeding the *MIV* of the raw transducers from being reached.

The burden is on the loudspeaker manufacturer to provide a different failure criteria than 3 dB of change to any one-third octave band if that is applicable to their loudspeaker. The tester and end user are not privy to the inner workings of the loudspeaker system, nor need they be. The loudspeaker is treated as a “black box” with an electrical voltage input and an acoustical voltage (pressure) output.

19.3.1 The Voltage Rating

It should be apparent that at the conclusion of the test that we know the maximum rms drive voltage that the loudspeaker can handle. This is all that a technician needs to know to avoid thermal damage to the loudspeaker when setting amplifier levels in the field. It is also all we need to know to select an appropriate power amplifier. The expression of this as a power rating is an optional step that can be muddied by how the loudspeaker’s impedance is determined, as well as marketing pressure to publish large power ratings. Loudspeaker manufacturers are encouraged to include the *MIV* along with power ratings on the specification sheet. This frees the technician from having to calculate the *MIV* from a (possibly exaggerated) power rating and ambiguous rated impedance.

19.3.2 Mechanical Limit Testing

In addition to thermal testing, it may be desirous to test the loudspeaker’s mechanical excursion limits. This test requires an impulsive stimulus in lieu of continuous noise. Don Keele has produced a peak tone-burst that is suitable for mechanical limit testing. The stimulus is a 6.5 cycle sinusoid in a cosine envelope at the desired test frequency. This short duration signal has a bandwidth of one-third octave. Three such signals spaced at one-third octave centers and mixed produce a one-octave test bandwidth. A duty cycle can be selected to emulate repetition rate of a kick drum used in electronic music. The output waveform of a measurement microphone is monitored on an oscilloscope as the drive voltage to the loudspeaker is increased. Visual deformation from the low voltage response indicates that the mechanical limits of the loudspeaker have been reached.

The *MIV-burst voltage* determined by this test can be very high. When used along with the rated impedance to calculate power, the result is a very high “burst power” rating for the loudspeaker. It is important to note that the actual power flow to the loudspeaker is fairly low due to the transient nature of the signal, which yields a very high crest factor. For example, if the *MIV-thermal* is 40 V_{rms} and the *MIV-burst* is 200 V_{rms} it is defensible to publish the loudspeaker’s voltage rating as follows:

1. MIV_{rms} = 40 V (IEC noise, 3-minutes).
2. MIV_{peak} = 200 V.
3. *Equivalent Amplifier Size* = 400 W (sine wave, 8 Ω).
4. *Burst MIV* = +14 dB.

These can be used to determine the power ratings by calculation using the rated impedance. I will describe the *Equivalent Amplifier Size* (*EAS*) later in this chapter. If the system designer specifies the amplifier based on the *EAS*, it will not have sufficient peak room to reach the *MIV-burst*, so there should be no danger of over-excursion. If the system designer specifies an amplifier that can pass the *MIV-burst*, then it will be capable of thermally destroying the loudspeaker. Which is used depends on many factors, including the application and the competence of the end user.

19.3.3 Conclusion

That’s the basic test. Some of the details, such as stimulus type, duration and failure criteria vary per standard. The method I’ve just demonstrated is just one of many possibilities. It is logical, repeatable, objective and easy to implement with readily available instrumentation. It should suffice to demonstrate the nature of power testing and loudspeaker power ratings.

19.4 The Amplifier as a Voltage Source

An amplifier has a maximum voltage amplitude that it can produce, as determined by its power supply rails. This establishes the highest peak that can pass through the amplifier. If excessive gain is applied to the signal, the waveform will be clipped as it tries to exceed the rail voltage. Most amplifiers provide a visual indicator for the onset of clipping.

The ideal amplifier is a constant voltage source. This means that the voltage of the audio waveform is not affected by the presence of, or changes to, the load impedance. As an example, let's fix the output voltage to 28 V_{rms} (about 100 W continuous into 8 Ω) and consider what happens as the load impedance changes, Fig. 19-4. An ideal (theoretical)constant voltage amplifier is an unlimited source of current. As the load impedance drops, the voltage is unchanged and the current increases, satisfying the power equation. Each halving of the load impedance doubles the output current and therefore the power into the load, so long as the voltage remains constant, which it can for this hypothetical amplifier. So, the more paralleled loudspeakers, the lower the load impedance, and the more current (and watts) from the amplifier. This makes it seem like a good thing to parallel more loads onto the amplifier, or to reduce the load impedance of the loudspeaker to a very low value to "get more watts," Fig. 19-5.

E_{sine}	I_{sine}	Z_{load}	*CAW*
28 V_{rms}	1.75 A	16 Ω	50 W
28 V_{rms}	3.5 A	8 Ω	100 W
28 V_{rms}	7 A	4 Ω	200 W
28 V_{rms}	14 A	2 Ω	400 W
28 V_{rms}	28 A	1 Ω	800 W
28 V_{rms}	56 A	0.5 Ω	1600 W
28 V_{rms}	112 A	0.25 Ω	3200 W

CAW = Continous Average Watts

Figure 19-4. Audio power from an ideal voltage source.

E_{sine}	I_{sine}	Z_{load}	*CAW*
28 V_{rms}	1.75 A	16 Ω	50 W
28 V_{rms}	3.5 A	8 Ω	100 W
24 V_{rms}	6 A	4 Ω	150 W
20 V_{rms}	10 A	2 Ω	200 W
?	?	1 Ω	?
?	?	0.5 Ω	?
?	?	0.25 Ω	?
	Not Recommended		

Figure 19-5. Audio power from a real-world voltage source.

A real-world amplifier is not an unlimited source of current, so this behavior no longer holds with decreasing load impedance. The voltage eventually drops as you daisy-chain additional loudspeakers onto the amplifier. The output power may continue to increase (the shaded region of Fig. 19- 6) but the voltage is dropping. Below about 4 Ω, this amplifier is no longer a constant voltage source. If these loudspeakers were spread along a parade route, this voltage drop will result in an audible drop in the L_P from the first loudspeaker as additional loudspeakers are daisy-chained along the route. So, adding more loads may increase the output power from the amplifier, but the sound level from each loudspeaker is reduced because the voltage is dropping. You get the highest L_P from the first loudspeaker if you disconnect the other ones. In short, the amplifier's best performance is achieved when it is operating as a constant voltage source, or stated another way, when it is not excessively loaded.

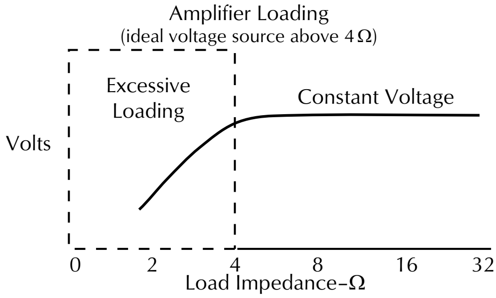


Figure 19-6. Real-world voltage source is current-limited.

19.5 The Equivalent Amplifier Size–*EAS*

If the *MIV* of the loudspeaker is known, we can calculate the required voltage capabilities of the power amplifier. The *MIV* was measured using a shaped-noise spectrum with 6 dB crest factor. This means that the peak voltage will be twice the rms voltage. For $MIV = 30\ V_{rms}$, the V_{peak} will be 60 V. Since amplifiers are rated using sine waves (3 dB crest factor), to pass a 60 V peak the sine wave rating must be 60 (0.707) = 42.4 V_{rms}.

The amplifier is a constant voltage source into 8 Ω, and an 8 Ω rating is typically given for a commercial amplifier, the sine wave power rating of the amplifier must be

$$\begin{aligned} W &= \frac{E^2}{R} \\ &= \frac{42.4^2}{8} \qquad (19\text{-}3) \\ &= 225\ \text{W} \end{aligned}$$

Eq. 19-3 calculates *EAS* from *MIV* (assumes 6 dB crest factor noise)

So, an amplifier with a sine wave power rating of 225 W will be able to pass the $MIV = 30\ V_{rms}$ to the loudspeaker for 6 dB crest factor noise. We can create a general equation for determining the *EAS* from the *MIV*. The *EAS* is the 8 Ω sine wave rating of the amplifier that can pass the test waveform at the conclusion of the *MIV* test.

Eq. 19-4 is the general equation to calculate *EAS* from *MIV*.

$$EAS = 10^{\frac{20\log MIV - 6}{10}} \tag{19-4}$$

The *EAS* assumes a 6 dB crest factor signal. It gives a practical amplifier rating for a given loudspeaker, assuming that it is desirous to achieve the maximum possible L_P without damage, which requires application of the *MIV* to the loudspeaker. The *EAS* gives a simple, one number specification for purchasing an appropriate amplifier for a given loudspeaker. It frees the designer from calculating the required amplifier size from the loudspeaker power rating, which is often exaggerated or ambiguous due to the impedance-dependence. It should be possible to drive this EAS amplifier to the onset of clipping without thermally damaging the loudspeaker. Of course, it is never a good idea to drive something to its thermal or mechanical limits, so operating the amplifier below clipping ensures a long life for the loudspeaker.

In contrast to a loudspeaker power rating, the *EAS* is an amplifier rating that is part of the loudspeaker's specifications.

Note that this is for 6 dB crest factor program material. In practice, it is likely that the crest factor will be higher. So, if the EAS-rated amplifier is driven to clipping with typical speech or music, the rms voltage to the loudspeaker should be below the *MIV* (due to the higher crest factor of real-world program material) and the loudspeaker should not be in danger of thermal damage.

The expression of the *MIV* as an *EAS* is only necessary because of the industry's insistence on rating amplifiers in terms of watts. If all loudspeakers were specified using the *MIV*, and amplifiers were rated in V_{rms} into a minimum Z (ideally expressed in dBV), power ratings could be avoided altogether. What matters is the voltage that the amplifier can produce across the load. Stated another way, for the amplifier to be a constant voltage source, it must be able to source the current demanded by the load impedance. If the load impedance is not too low, its current demands can be satisfied by the amplifier and need not be considered further. This allows the signal transfer between amplifier and loudspeaker to be assessed by voltage only. If not, then the combination of voltage and impedance, or voltage and current, must be considered. Power is just a way to express this as a single number. Eq. 19-5 are power equations.

$$\begin{aligned} W &= \frac{E^2}{R} \\ &= IE \\ &= I^2R \end{aligned} \tag{19-5}$$

where,
W is the power in watts,
E is the electro-motive force in volts,
I is the current in amperes,
R is the resistance in ohms.

A power rating alone is useless, because it doesn't yield sufficient information for determining the amplifier's output voltage. A "100 W" amplifier could be 33 V @ 3 A, 3 V @ 33 A, 100 V @1 A, or 1 V @ 100 A.

19.6 Power from a Voltage Source

Let's move past the practicality, simplicity and elegance of considering only the voltage from the amplifier. This can be necessary when handling exceptions and special cases. The ability of the amplifier to maintain its output voltage depends on its ability to source the current required by the load. If the current demand gets too high, the voltage may sag. Let's examine some factors that tax the current production of the amplifier.

Amplifiers don't make power—they convert it. The audio power that comes from an amplifier must in turn come from the utility outlet that it is connected to. This is why the wire gauge of the line cord matters, and why long extension cords can be a problem due to voltage drop across the wire's resistance (line loss). The maximum output voltage from an amplifier can be determined by measuring it into an open circuit, which is a load with a very high impedance relative to the amplifier's source impedance. For modern amplifiers, this typically means 8 Ω or higher, given the fractional output impedance (0.01 Ω or less) of modern amplifiers.

With a resistive load present, the power flow can be determined by the power equations previously given in Eq. 19-5.

This is the "apparent power" from the source. Apparent power (in volt-amperes) assumes that the voltage and current are in phase, which is only the case for a purely resistive load. The actual power in watts is likely less than the apparent power given the nature of real-world loads, where reactive (storage) properties reflect some of the power back to the amplifier from the load. In depth calculations account for this by considering the phase angle between the voltage and impedance, or a "power factor." I'll assume a resistive load for now and use watts as the unit. The power equation says that the apparent power from a 120 V_{rms}, 20 A utility power circuit is 2,400 W. If an amplifier were 100% efficient, it could produce 2,400 W or 34 dBW of audio power from the 120 V, 20 A outlet when passing a sine wave. The conversion efficiency of modern switch-mode amplifiers–Class D and its variants–can exceed 80%, meaning that an amplifier with a sine wave rating of 2 kW continuous is a possibility. In effect, the largest amplifiers can produce sine waves that are very close in voltage, current and therefore power to a household electrical circuit in the USA—an incredible achievement in amplifier evolution. Of course utility power is a 60 Hz sine wave. Music and speech signals are broad band, and will produce considerably less power than a pure sine wave due to their higher crest factor. Let's go with this theoretical 2,400 W amplifier (34 dBW) for now. How much audio power can you really expect from this amplifier?

The amplifier's job is to accurately reproduce the signal voltage presented to its input. It's not technically correct to call it a power amplifier, because the output power has nothing to do with the input power—only the input voltage. It has to amplify this voltage and preserve the wave shape while sourcing the current being demanded by the load. As I showed earlier, this gets tougher as the load impedance drops. For a constant voltage interface (any modern power amplifier), current, and therefore power, is a responder. It is the "result" of the drive voltage and load impedance. Since the load impedance is fixed (unless your loudspeaker overheats, causing its impedance to rise), the current (and the resultant power) can only be changed by adjusting the drive voltage to the amplifier. For this reason, I always do my initial assessment of an amplifier by measuring its maximum open circuit sine wave output voltage, followed by verifying that this voltage can be maintained into an 8 Ω resistor. Next I consider the minimum load impedance for which the amplifier can sustain this voltage, using a1 dB drop as the failure criteria. The maximum output voltage can be measured with an rms voltmeter or oscilloscope. It can be determined by calculation from the 8 Ω power rating provided by the manufacturer using the power equation, solved for voltage.

$$E = \sqrt{WR} \qquad (19\text{-}6)$$

where,

W is the 8 Ω power rating,

R is the load resistance.

It is normal to find excellent correlation with the amplifier's published specifications. I looked at several of the largest available amplifiers and assessed their voltage output. An increasing number of them have approximate maximum sine wave voltages in the 130 V_{rms} range, which is a peak voltage close to 200 V, found by multiplying V_{rms} by 1.414. The term "200 volt rails" is gaining popularity among the amplifier crowd. A sine wave of 140 V_{rms} will produce the following wattages into a test resistor from an ideal amplifier (one that can source the required current):

1. 2,500 W into 8 Ω.
2. 5,000 W into 4 Ω.
3. 10,000 W into 2 Ω.*

*theoretical value only!

Note that all of these rated powers exceed the 2400 W that is available from the electrical outlet, especially if the amplifier has multiple channels. At these wattages, the available utility power may be the limiting factor with regard to what amplifier can output. To assess the music/speech power, I'll take the rail voltage and convert to dB, and subtract the crest factor of the program material. I'll be generous and assume a crest factor of 10 dB, which is typical of slightly clipped pink noise, in other words, an amplifier being driven pretty hard. I'll then convert the level back to voltage, the field quantity upon which all power ratings are based.

$$dBV_{peak} = 20\log 200$$
$$= 46 \text{ dBV}$$

$$dBV_{avg} = 46 \text{ dBV} - 10 \text{ dB}$$
$$= 36 \text{ dBV}$$

$$E_{max} = 10^{\frac{36}{20}}$$
$$= 63 \text{ V}_{\text{rms}}$$

When applied to an 8 Ω resistive load the power is:

$$W = \frac{63^2}{8}$$
$$= 496 \text{ W}$$

This is a far cry from the sine wave power rating of the amplifier, and well within the ability of the utility power circuit.

If this amplifier can sustain 63 V_{rms} into a 4 Ω load (it probably can), the power doubles to

$$
\begin{aligned}
W &= \frac{63^2}{4} \\
&= 992\ \mathrm{W}
\end{aligned}
$$

As you can see, the power drops dramatically when you substitute real-world program material for the sine wave. Now, keep in mind that sometimes real world program material is a sine wave. A synthesizer can produce very low crest factor signals that can trip circuit breakers and burn-up a subwoofer connected to the amplifier. But, this is the exception, not the rule. The vast majority of loudspeakers in the marketplace have rated impedances of 4 Ω or higher. The average impedance is higher yet, which is what the amplifier sees if you are playing broadband program material (speech or music). So, to say that we can draw 1 kW from a "2400 W" amplifier is being generous.

It would not be unusual to be able to drive our 2,400 W amplifier to clipping into a modern line array box using a pink noise signal. Does that mean that the box is "handling" 2,400 W? No. The high crest factor of the pink noise and the real-world higher-than-rated average impedance of the loudspeaker means that the power flow is far less than 2,400 W. It is handling the full voltage swing of the amplifier, but not its full output power. The high crest factor signal and real world loudspeaker impedance can allow me to access the full voltage swing of a theoretical amplifier rated at 5 kW/channel into 4 Ω. But, the amplifier will run out of voltage swing before it runs out of power.

The impedance of a real-world loudspeaker is higher than its "rated" impedance over most of its bandwidth, Fig. 19-7. While some amplifiers can drive "2 Ω loads" it would be rare to ever encounter a true 2 Ω load in practice other than with a bank of test resistors. Besides the voltage drop, there are other reasons why very low impedance loads (i.e. 2 Ω) are problematic, including the requirement for very heavy loudspeaker cable, and the detrimental affect on damping factor. While you may "get the most watts" into a 2 Ω load, the amplifier will perform better if it is not so heavily loaded. We wouldn't dream of putting weights in the trunk of our car to make it produce more power to attain a target speed. It is equally silly to load down an amplifier for the sole purpose of "getting more watts."

19.7 Burst Testing

If the power from a source cannot be sustained over time, it may be interrupted periodically to avoid damage to the source or the production of excessive heat in the load. Examples include pumping the brakes on your car when descending a steep hill, or breaking a long weld into shorter segments. A source that cannot run continuously without overheating may be given a duty cycle rating. Playing a repetitive waveform (like a kick drum) is like pumping the brakes on your car. The power and heat production are reduced when compared to the steady-state condition. That's a good thing, since otherwise we'd be tripping circuit breakers and burning voice coils. Ironically, the largest power ratings on an amplifier spec sheet result from "Burst testing." This is a method developed for measuring cell phone signals, and it is also used on power amplifiers. A burst test hits the amplifier with a signal that lasts only a few milliseconds, Fig. 19-8. These bursts produce some very high amplitude peaks, but only for a very short period of time. Looked at another way, burst testing can prevent the circuit breaker from tripping (a good thing) and allow our 2,400 W amplifier to drive a 2 Ω load at its full output voltage. This theoretically doubles the power to nearly 5 kW from a single amplifier channel. Is this a legitimate power rating for the amplifier?

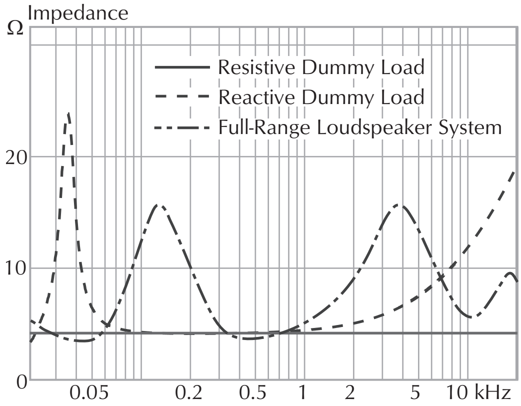


Figure 19-7. Load impedance seen by amplifier.

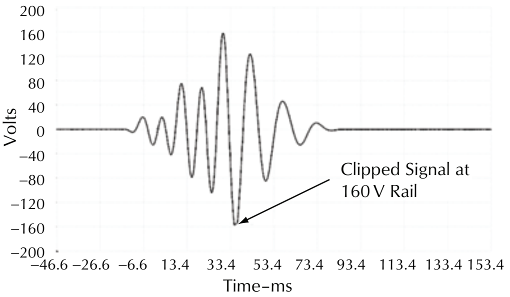


Figure 19-8. 6.5 cycle tone burst for 1⁄1 octave band centered at 80 Hz.

The moment-by-moment voltage amplitude values of an audio waveform can be very high. Multiplying the instantaneous voltage by the instantaneous current yields the instantaneous power. Instantaneous power is a means to an end, not the end itself. It must be integrated overtime to find the root-mean-square voltage and the average power. While marketing types love to publish burst power specifications, and end users will make buying decisions based on them, it is ill-advised to allow characteristics that are transient and difficult to assess (i.e. peaks) to strongly influence the buying decision.

While burst signals can have very high amplitudes, they are quite short in duration. The waveform produced by a kick drum is made up of multiple frequencies that are not phase coherent in the load due to the non-linear phase response of the subwoofer, which increases the crest factor of the waveform. The audio power produced by a kick drum is not all that high, due to its high crest factor and low duty cycle. A far more punishing waveform would be the continuous sine or square wave output from a synthesizer. This is the only way that you will challenge an amplifier to produce its full rated power, and you may burn up your subwoofer in the process. While it is a good thing for an amplifier to be able to produce waveforms with very high instantaneous power, it is not often clear to consumers what this means and under what conditions it is measured (the fine print). All they see is a huge number with "watts" after it and may assume that it is a continuous rating and consider it as such. Or worse yet, they may think that it is something they need and something that the loudspeaker has to handle.

19.8 Power Rating Possibilities

There are many possible tests for rating the power output of an amplifier. The most brutal is the use of a continuous sine wave, with the amplifier connected to a resistive load and operated for a specified period of time—"continuous" implies "indefinitely."

I am a big fan of this method, for several reasons.

1. It's a simple test. There's no way to fudge it, fool it, fake it or misrepresent it.
2. The rating can be easily verified in the field with simple instrumentation.
3. I can easily determine the output power with other waveforms (music or speech) by substituting the crest factor of the waveform for that of the sine wave. This always results in less power flow due to the higher crest factor of the more complex waveforms.
4. The distortion of a sine wave is much easier to measure than the distortion of more complicated waveforms, so it is easy to determine when the amplifier has reached its maximum linear output level.

As a system designer, I want a simple, no-nonsense conservative rating of what the amplifier can do with a known waveform (i.e. a sine wave). I can increase the crest factor (de-rate the amplifier) to know the level that will be produced by music or speech. Note that it is unlikely that the amplifier will ever have to pass a full-scale sine wave into a load for an extended period of time in an audio application, but the continuous sine wave power rating is still a useful rating method and benchmark.

While the term "power amplifier" enjoys widespread use, it has led to a "power mind set" when interfacing amplifiers to loudspeakers. This is unfortunate, and has contributed to many misconceptions in the marketplace. The frequency response of a loudspeaker is a function of the drive voltage to the loudspeaker, not the applied power. The voltage delivered to the loudspeaker by the amplifier is independent of the loudspeaker's impedance. I can sweep a $2.83\,V_{rms}$ sine wave from 20 Hz to 20 kHz and the loudspeaker gets exactly $2.83\,V_{rms}$ at every frequency. The current drawn from the amplifier, and therefore the power, varies as a function of frequency due to the complex impedance curve of the loudspeaker. An attempt to feed a loudspeaker the same power in each $1/N$-octave band would result in a very uneven acoustical frequency response from the loudspeaker.

Adjustments to this drive voltage are directly observable in the loudspeaker's frequency response. Loudspeaker equalization is accomplished by modifying the drive voltage, not the applied power. It is proper to consider the amplifier as a voltage source with sufficient available current to satisfy the demands of the loudspeaker's impedance. While the product of voltage and current can be expressed as power in watts, keeping them separate yields much more insight into the amplifier/loudspeaker interface. As I have shown, it is the applied voltage to the loudspeaker which is the primary quantity of interest.

19.8.1 Real World Power Generation

To summarize, it is theoretically possible for a highly efficient audio power amplifier to produce continuous sine wave power in the 2 kW range when connected to a 2400 W utility circuit. For it to actually have to do this is unlikely, given the nature of music and speech signals and the real-world impedance of loudspeakers. Rating the amplifier based on a very low load impedance, tone burst signal and in linear units ("watts") is definitely putting the best face on its performance. Since no standard exists for "burst testing" amplifiers, you can bet that it is done differently by every manufacturer, since the burst duration and duty cycle affect the results. This doesn't make the numbers meaningless, but you won't be able to use them to compare different makes and models.

It would be better to say that this amplifier has a rail voltage of 200 V for passing the peaks in the audio waveform, and that it can sustain a sine wave with this peak value indefinitely into an 8 Ω load, and for a few cycles into lower load impedances. All of this can be determined from a simple 8 Ω sine wave power rating.

19.8.2 Things Amplifiers Hate

There are two conditions that power amplifiers don't like—low crest factor signals and 2 Ω loads. The continuous sine wave testing is by far the most revealing with regard to how much power the amplifier (and the electrical outlet) can source. Few if any power amplifiers can maintain their output voltage into 2 Ω with a sine wave source. This makes the 8 Ω sine wave rating attractive to the system designer as a guaranteed performance measure.

19.8.3 Burst Testing Results

Amplifiers get their largest power ratings from tone burst testing. Ironically, that is the easiest signal for most amplifiers to pass, allowing them to produce their maximum voltage for a few tens of milliseconds, even into 2 Ω with both channels driven. It's a shame that this number wields so much influence in the marketplace.

19.8.4 Conclusion

Some general conclusions regarding amplifier power ratings can now be drawn. Here are some of the more important ones.

1. A 1 kHz sine wave rating into 8 Ω is the best measure of an amplifier's performance for a sound system designer. The amplifier acts as a constant voltage source into 8 Ω, and its output voltage is typically independent of frequency over the vast majority of the amplifier's pass band. The sine wave voltage can be scaled to any crest factor by calculation. This makes amplifier selection and accurate sound pressure level calculations at the drawing board possible.
2. Don't load your amplifiers to 2 Ω to "get more watts." When the voltage sags, even if the power increases, you are losing output level and engaging protection mechanisms.
3. Amplifiers don't like low crest factor signals. If you excessively compress or limit the program material, and drive the amplifier to clipping, you are likely engaging the amplifier's protection mechanisms, with potentially audible ramifications.
4. Don't compare amplifiers using power ratings derived from burst testing. These are vanity specs that look impressive but reveal very little about the amplifier's performance.

The sound system designer is well-served by the simplest amplifier testing and rating method. Sine waves are commonly used to establish the maximum output level of the other constant voltage sources in the signal chain, such as mixers and signal processors. Sine wave amplifier ratings, ideally expressed in dBW rather than watts, provide a simple, logical way of rating power amplifiers.

19.9 Putting It All Together

Part of the commissioning process for any sound system is the establishment of the L_P at the listener's ears. Here is a step-by-step approach.

With the mixer producing an average level of +4 dBu for broad band program material, pink noise or music, and the signal processing appropriately configured and set at or near unity, a line level signal should be present at the amplifier's input. The amplifier's sensitivity control determines the output voltage produced by the amplifier in response to this input voltage. This ultimately determines the electrical power applied to the loudspeaker, as well as the L_P produced from the loudspeaker.

19.9.1 A Warning

Before proceeding, a word of warning. Improper setting of the amplifier's sensitivity can damage loudspeakers and hearing. Before turning on amplifiers be sure that:

1. Any signal processing required by the loudspeaker is in place and properly configured, especially crossover networks.

2. The amplifier(s) is properly connected to the loudspeaker. Be especially careful with biamped and tri-amped systems, as there are no standards that govern the wiring of multi-pin connectors.
3. The amplifier(s) is appropriate for the loudspeaker, based on the *MIV* described previously.
4. The L_P is properly monitored during adjustment and kept at safe levels.

19.9.2 A Common Misconception

The amplifier's input sensitivity does not affect nor determine the maximum voltage that can be produced by the amplifier or the power rating of the amplifier. Some users feel that unless the input sensitivity is maximized, they can't get all of the power. Not true. This control determines the input voltage required to produce a given output voltage. If you turn it down, you can still drive the amplifier to clipping. It just takes more input voltage to do so.

19.9.3 The Input Sensitivity Control

The input sensitivity control may be on the front or back of the amplifier. One can make an argument for either. Consider the extreme settings of this control.

Fully counter-clockwise is the minimum sensitivity setting. On most amplifiers, this is"off." Simple enough.

Fully clockwise is the maximum sensitivity setting. This produces the highest output voltage from a given input voltage. Many use this as the default setting, but there can be ramifications.

Some amplifiers have selector switches that scale the sensitivity range.

There exists several philosophies for setting the amplifier's sensitivity, justified by the various applications for which the amplifier may be used. I'll consider the major ones here.

Before I get into the details, I'll start by saying that for quick setups and for simple systems, the final step of system gain structure is to simply power up the amplifier and increase input sensitivity until the desired L_P is produced. For many systems, it really is that easy. Just make sure that there is a strong meter reading on the mixer's meter (i.e. 0 VU) so that you have a visual level reference while mixing.

This is the correct sensitivity setting so long as the amplifier isn't clipping, the L_P isn't dangerous, and the loudspeaker isn't overheated. Don't let the following discussion of the details confuse you. If done in the proper order, setting the amplifier's sensitivity is no more difficult than adjusting the volume of your radio or TV.

Doing it in reverse order by starting with the amplifier at maximum sensitivity can make it much harder, or even impossible, to achieve a good system gain structure. In some cases we may end up with it fully clock-wise, but we don't want to start at maximum sensitivity.

The Details

At their maximum sensitivity setting, most amplifiers can be driven to full output voltage by just over one volt at the input. A common example is a sine wave of +4 dBu or 1.23 V_{rms} that falls within the rated bandwidth of the amplifier. It may be a bit higher or lower, depending on the make and model.

At full operating level the mixer should be producing about +4 dBu average, plus peaks of +20 dB, for typical program material. This is a maximum output level of about +24 dBu. The signal processing is passing this level at unity. This means that the signal level at the amplifier input is about +20 dB relative to what the amplifier needs to produce its full output. That's fine, as it simply means we have more voltage than what we need to access the amplifier's full voltage swing. The amplifier's sensitivity can be reduced accordingly. It's usually better to have too much of something than not enough. By giving the amplifier a higher sensitivity than is needed, manufacturers assure broader compatibility. Input sensitivity can be reduced by the simple adjustment of a control.

In a modern system, the signal level presented to an amplifier may be far lower than +4 dBu. The most common cause is the integration of consumer gear into professional systems. It's also becoming increasingly common for digital signal processors (DSP) to introduce 10 dB or more of attenuation to the signal rather than operating at unity. This is one way that they can handle the high voltages produced by some mixers. If a crossover network is part of the DSP configuration, the dividing of the spectrum reduces the signal level.

Amplifiers need adequate "gain reach" to handle any of these situations.

Setting the Amplifier's Sensitivity

Amplifiers should always be set using a broadband signal. Pink noise works especially well, as it emulates an intense music signal in both spectral content and crest factor. Unlike music or speech, it is easily measured with a digital voltmeter.

Ultimately, the amplifier's sensitivity setting determines the L_P produced at the listener.

That becomes the main consideration for the correct setting. When the L_P is right, the sensitivity is right. Of course the L_P should be appropriate for the application and not pose a danger to listeners. Here's an orderly, logical method for setting the control.

Step 1 is to set the amplifier sensitivity to the fully counter-clockwise position. The mixer should be producing about 1 V_{rms} of pink noise. The signal processing should be in place and configured correctly for the loudspeaker attached to the amplifier.

Advance the input sensitivity slowly until:

1. It's loud enough, or
2. The amplifier clips, or
3. The *MIV* of the loudspeaker is reached.

Hopefully Item 1 is realized before Items 2 or 3. Let's consider some caveats of each.

1. It's loud enough. For many systems this can be accomplished by just listening to some familiar program material, ideally similar to what the system will be used for. This empirical approach is the fastest way to set the sensitivity. You may want to involve the client, and if you're brave, a committee.

   Simple listening can be used if there is no danger of clipping the amplifier or overpowering the loudspeaker. It's always interesting to get a measured L_P on what you have determined by listening, for documentation.

   Pink noise and a sound level meter may be required if the system has a specified L_P that it must produce. Always make sure that this is a realistic target that will not expose listeners to damaging levels.
2. The amplifier clips. The clip light on the amplifier lets you know when you have hit the power supply rails with the signal peaks. Unlike line level devices, amplifier lights usually illuminate at actual clipping. Slight clipping is usually not audible and poses no threat to loudspeakers if the peaks are within the loudspeaker's excursion limits.

   In any event, when the clip light comes on, you must stop increasing the sensitivity. If the L_P must be higher, you'll need a larger amplifier or a more efficient loudspeaker.
3. The *MIV* of the loudspeaker is reached. A true rms voltmeter can be used to monitor the amplifier voltage as the sensitivity is increased. If you reach the *MIV* and it's still not loud enough, you should assess the situation carefully. Higher L_P may be ill-advised.

Fig. 19-9 shows the voltmeter connection and "stop criteria" for setting the amplifier's sensitivity.

19.9.4 The 30/30 Guideline

Many systems can be "roughed in" using the 30/30 guideline. Here's how it works:

- At 30 ft from the loudspeaker, about 10 m, the sound level will be about −20 dB relative to the level at 1 m, due to inverse-square law level change.

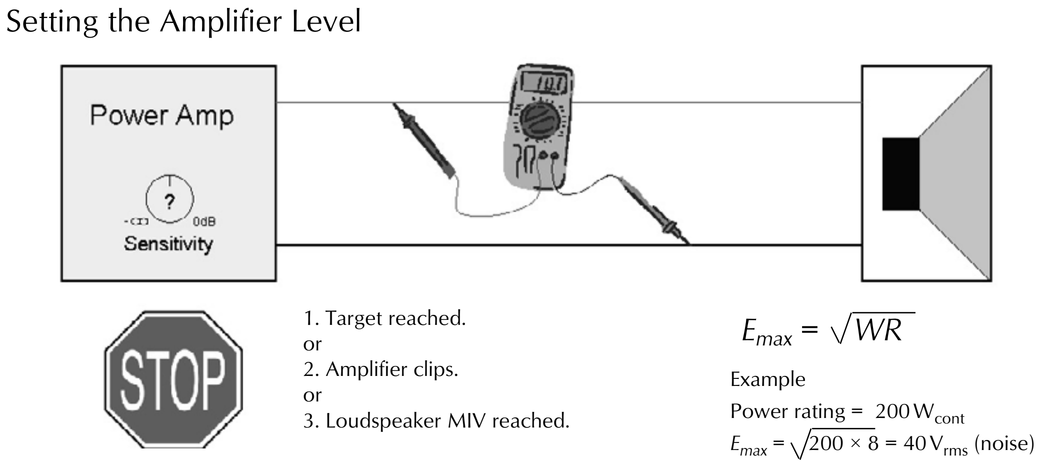

Setting the Amplifier Level
Power Amp
?
-∞
0dB
Sensitivity
STOP
1. Target reached.
or
2. Amplifier clips.
or
3. Loudspeaker MIV reached.
$E_{max} = \sqrt{WR}$
Example
Power rating = 200 W_{cont}
$E_{max} = \sqrt{200 \times 8} = 40 V_{rms}$ (noise)

Figure 19-9. Criteria for setting the amplifier's sensitivity.

$$dB = 20\log(3.28/30)$$
$$= -20 \text{ dB}$$

- At 30 V_{rms} applied to the loudspeaker, the power is approximately 100 W continuous into an 8 Ω loudspeaker. This is about +20 dB relative to the 2.83 V_{rms} used to measure the loudspeakers sensitivity. Also, this is well within the power handling capabilities of most full-range medium to large format sound reinforcement loudspeakers.

$$dB = 20\log(30/3.28)$$
$$= 20 \text{ dB}$$

- The average sensitivity of the loudspeaker can be read from its specification sheet or data viewer, Fig. 19-10.

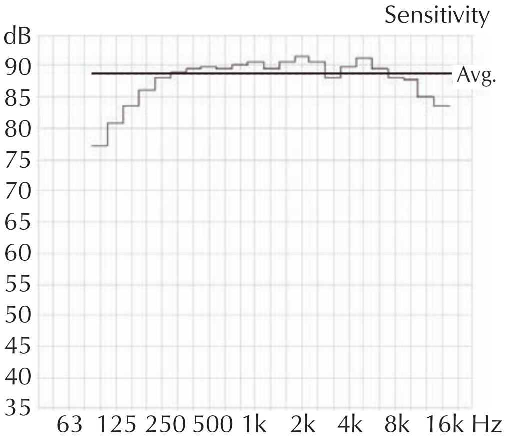


Figure 19-10. Average loudspeaker sensitivity as shown in the Common Loudspeaker File format viewer. (*Courtesy www.clfgroup.org.*)

This means that if you adjust the amplifier input sensitivity to produce the loudspeaker's average sensitivity at 30 ft, you are feeding about 100 W continuous to the loudspeaker. This is a good place to pause when setting the playback level of the system. You can then assess the situation.

Of course, you can use any target L_P. The loudspeaker's average sensitivity is just a guideline. If the system is loud enough at a lower L_P, stop there.

- Is it loud enough? If not how much more level is needed? Every additional 3 dB doubles the power to the loudspeaker, and the required size of the amplifier. Proceed with caution!
- Is the amplifier clipping? If it is, reduce the sensitivity.
- Is the loudspeaker strained? If the sonic character of the noise changes as you approach this level, turn it down!

A Word of Warning

An amplifier set to maximum sensitivity can be a disaster waiting to happen. Any applied signal will be passed on to the loudspeaker at a potentially very high amplitude, including pops, glitches, feedback and other undesirables. The sensitivity control is the last line of defense against such signals, so be VERY careful when max-ing out amplifiers—especially big ones!

Some Practical Considerations

While it is technically correct and in most cases advisable to operate amplifiers at less than their maximum sensitivity, for some types of systems this can present problems. Consider the following:

1. Exposed sensitivity controls can be tampered with. A passer-by or ill-informed user may max-out your carefully adjusted sensitivity knob and think they are doing you a favor.
2. Systems that have many amplifiers, such as arenas and touring rigs, are usually controlled by a computer network. If the physical sensitivity control is reduced, the full control range may not be accessible from the PC. In either case, it is a common practice to operate amplifiers with their sensitivity control at its maximum setting. The level is attenuated by software or by turning down the output stage of the device directly preceding the amplifier, usually a digital signal processor (DSP). The down side is that some amplifiers are noisy with their sensitivity set at maximum.

You can easily arrive at the required attenuation ahead of the amplifier to allow maximizing the amplifier's sensitivity if the control is marked in dB. "Gain trading" transfers the "dB below maximum" from the amplifier's input sensitivity control, to the output of the previous stage, Fig. 19-11.

19.10 Multi-way Loudspeakers

Some loudspeakers require more than one amplifier. For tri-amped systems start by setting the sensitivity for the mid-frequency component for the target L_P. Advance the LF and HF sections to balance with MF.

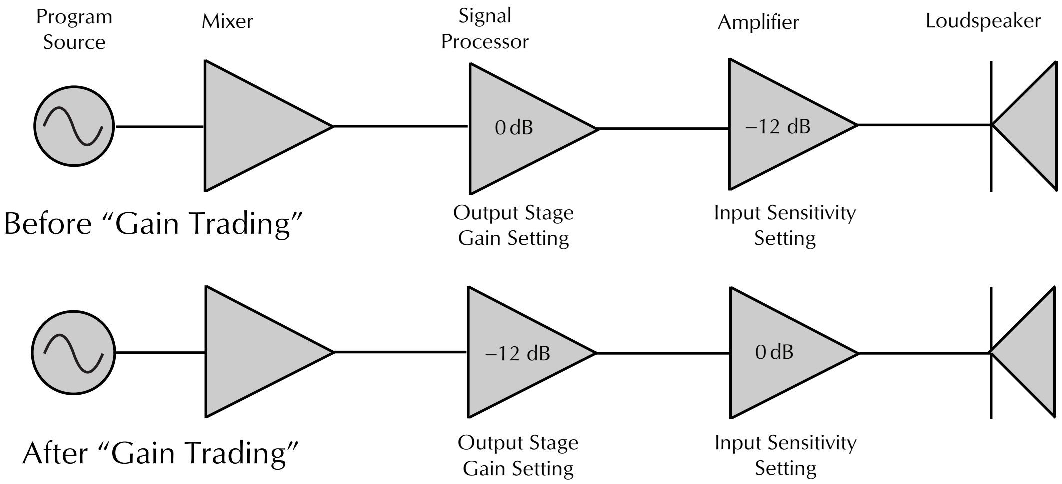


Figure 19-11. Gain trading to maximize amplifier input sensitivity

This can be done by listening, or better yet, with a ⅓ octave real-time spectrum analyzer.

For bi-amped systems, start with the LF section and then add HF for the proper balance.

19.10.1 Fixed Gain Settings

Some amplifiers allow the voltage gain to be fixed to a certain value, such as 26 dB or 32 dB. The input sensitivity control is then operated at maximum. This can facilitate driving multi-way loudspeakers from DSPs, where the spectral balance set in the DSP needs to be preserved through the various sized amplifiers that drive the loudspeaker system. It can also facilitate mixing multiple brands of amplifiers. In this application the amplifier just becomes a voltage gain block in between the amplifier and loudspeaker.

A specific amplifier gain setting may be required for the limiter stages in the DSP to function correctly for a given loudspeaker model. Consult the DSP manufacturer for the proper setting of these switches.

Note that these don't change the maximum possible output voltage of the amplifier, only the input voltage required to reach it.

19.10.2 Powered Loudspeakers

Some loudspeakers have internal amplifiers. If a sensitivity control is provided, it is adjusted exactly as if the amplifier were separate from the loudspeaker.

If no control is provided, then the level is set in the last stage of the signal processor driving the loudspeaker. This allows the mixer to be operated in its proper level range (i.e. 0 VU) without producing excessive L_P in the audience. If no signal processor is present, add one. In some cases it is necessary to externally attenuate the mixer's signal if it is directly connected to a powered loudspeaker that has no sensitivity control.

19.11 System Gain Structure

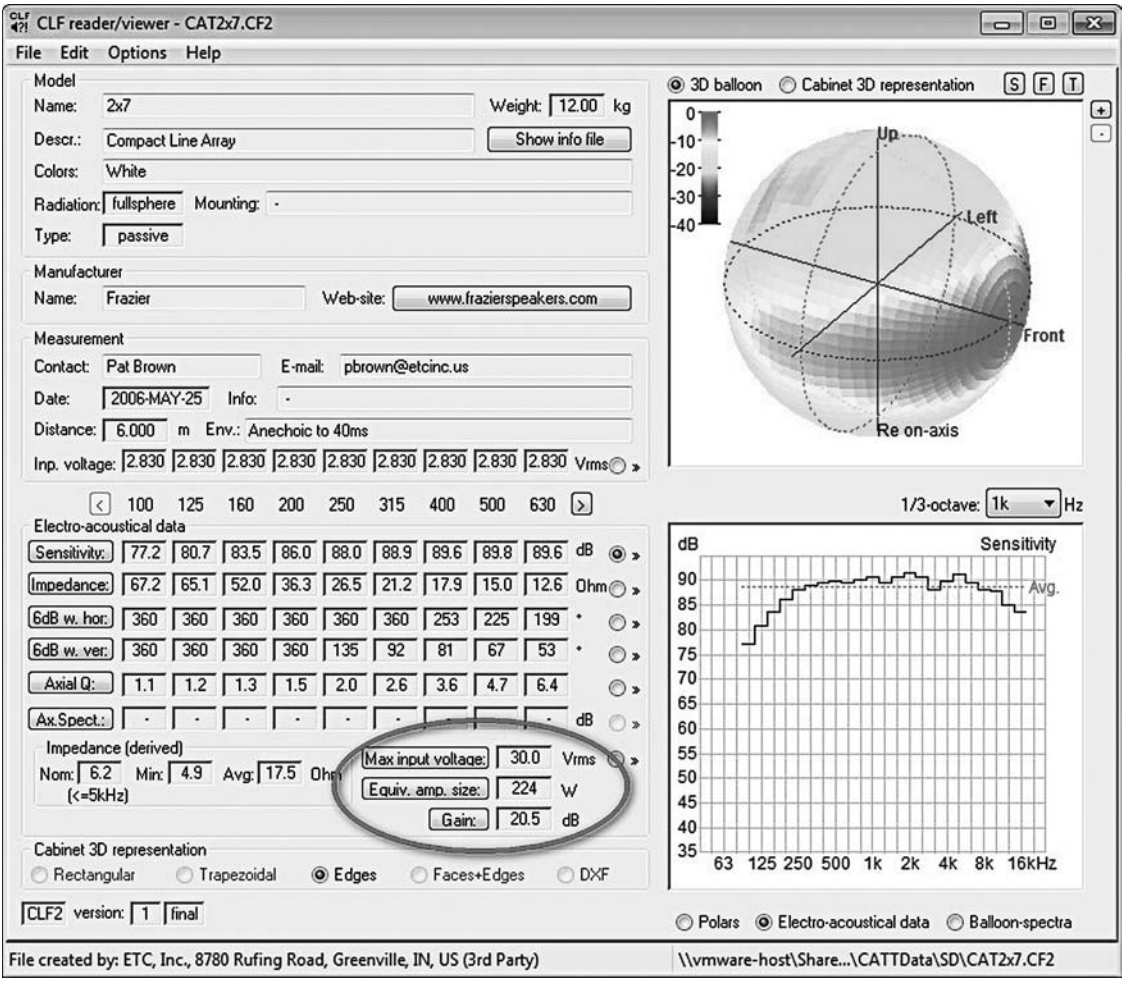
The best system gain structure results when each component in the signal chain is operating in the optimum part of its dynamic range. Mixers and signal processors incorporate metering to provide visual indication that this has been established. The amplifier's job is to produce the gain needed to achieve the desired L_P at the listener. The amplifier's input sensitivity is variable for this reason. It's setting is correct when the desired L_P is achieved in the house with the mixer's meter indicating that it is operating with good signal-to-noise ratio and sufficient peak room. Historically this is 0 VU on a volume indicator, or an average level of −20 dB relative to the full output voltage of the mixer.

19.12 Combining *MIV* and *EAS*

The CLF data viewer gives all of the information needed by the sound system designer for selecting a loudspeaker and amplifier. Fig. 19-12 shows the freeware CLF Viewer.

Note that this loudspeaker has an average sensitivity of 88 dB, and an *MIV* of 30 V_{rms}. The *EAS* is 224 W, so if an amplifier with a sine wave rating of this value is used with the loudspeaker, about 20 dB of gain above the average sensitivity can be expected at 1 m. This is for 6 dB crest factor program. The L_P can be reduced accordingly for higher crest factor signals. This amplifier, if driven to the threshold of clipping with typical program material, should pose no thermal danger to the loudspeaker. The 1 m maximum L_P can be extrapolated to any distance using the inverse-square law.

Most computer room modeling programs can utilize this data file directly for predicting the performance of a system.



CLF reader/viewer - CAT2x7.CF2

Model

Name: 2x7

Weight: 12.00 kg

Descr.: Compact Line Array

Colors: White

Radiation: fullsphere

Mounting: -

Type: passive

Manufacturer

Name: Frazier

Web-site: www.frazierspeakers.com

Measurement

Contact: Pat Brown

E-mail: pbrown@etcinc.us

Date: 2006-MAY-25

Info: -

Distance: 6.000 m

Env.: Anechoic to 40ms

Electro-acoustical data

	100	125	160	200	250	315	400	500	630	
Inp. voltage:	2.830	2.830	2.830	2.830	2.830	2.830	2.830	2.830	2.830	Vrms
Sensitivity:	77.2	80.7	83.5	86.0	88.0	88.9	89.6	89.8	89.6	dB
Impedance:	67.2	65.1	52.0	36.3	26.5	21.2	17.9	15.0	12.6	Ohm
6dB w. hor:	360	360	360	360	360	360	253	225	199	°
6dB w. ver:	360	360	360	360	135	92	81	67	53	°
Axial Q:	1.1	1.2	1.3	1.5	2.0	2.6	3.6	4.7	6.4	
Ax.Spect.:	-	-	-	-	-	-	-	-	-	dB

Impedance (derived)

Nom: 6.2 (<=5kHz)

Min: 4.9

Avg: 17.5 Ohm

Max input voltage: 30.0 Vrms

Equiv. amp. size: 224 W

Gain: 20.5 dB

Cabinet 3D representation

○ Rectangular ○ Trapezoidal ● Edges ○ Faces+Edges ○ DXF

CLF2 version: 1 final

● 3D balloon ○ Cabinet 3D representation

1/3-octave: 1k Hz

○ Polars ● Electro-acoustical data ○ Balloon-spectra

File created by: ETC, Inc., 8780 Rufing Road, Greenville, IN, US (3rd Party)

\\vmware-host\Share...\CATTData\SD\CAT2x7.CF2

Figure 19-12. The CLF Viewer. (*Courtesy www.clfgroup.org.*)

Chapter 20

Computer-Aided System Design

by Pat Brown

20.1 Spherical Loudspeaker Data 443
20.2 Near Field vs. Far Field 444
20.3 The Measurement Process 445
20.4 Loudspeaker Arrays 447
  The Dynamic Link Library—DLL 447
  Magnitude vs. Phase Data 448
20.5 Direct Field Modeling 449
20.6 Room Model Detail 449
  Predicting Room Reflections 450
  The Objective of Room Modeling 451
20.7 Room Acoustics—An Overview 451
  The Room Impulse Response—*RIR* 451
  Post-Processing the *RIR* 452
20.8 Absorption 453
  Modeling with Absorption Coefficients 453
  Sound Behavior 454
  The Direct Sound Field 454
  The Early-Reflected Sound Field 455
  Late Reflections or Echoes 455
  Reverberation 456
  Room Measures 457
20.9 Realistic Room Models 457
20.10 Universal Room Modeling Tips 460
20.11 Conclusions 461
References 461

My first foray into sound system design was in the late 1970's. I had installed a sound system in an octagonal worship space, designed for pipe organ support. The poor speech intelligibility of my design was obvious to all, including me. It was then that I found the first edition of this book, *Sound System Engineering* by Don and Carolyn Davis at the local electronics parts house. The now famous "yellow book" consumed much of my time from that point on, clearly showing that I was making some serious mistakes in my system designs. Every principle learned has proven timeless. What we have gained in the last 40 years is the ability to crunch the numbers faster and display the prediction results with stunning graphics. But, the underlying principles have remained unchanged. Any serious designer must begin with those. Only then should one consider using software tools to speed up the process. Before using a computer, one must learn what to compute. It is truly humbling to contribute to the fourth edition of the book that has formed my career.

I'll begin this chapter with a surprising statement—"Computers can't design sound systems." While this may not be true in the future, it is true now. Many potential sound system designers have been dismayed after purchasing a sophisticated room modeling program, only to find out that it is basically a calculator that executes algorithms that are based on approximations and assumptions regarding sound wave behavior. They are further dismayed to find that the usefulness of the program is limited by their own understanding of basic acoustics and electro-acoustics. That is not to say that these programs aren't sophisticated. Some are incredibly complex, even artistic—a mixture of deterministic calculations and software-specific algorithms for simulating acoustic behavior based on geometric principles. But, there will always be a difference between acoustic predictions and reality.

Do sound waves behave like light rays? "Yes," in some ways, but "no" in others. A failure to realize this will lead to surprises.

For an excellent overview of the history of room acoustics modeling, see *The Early History of Ray Tracing in Room Acoustics* by Peter Svensson. This work clearly shows that acoustics modeling is a game of approximations, assumptions and compromise. It's not simply a matter of number crunching. Faster processor speeds and breathtaking graphics have not removed the fundamental limitations of modeling sound geometrically based on optical principles, known as Geometrical Acoustics or GA.

That said, I am an enthusiastic supporter and advocate of computerized room modeling as a tool for the sound system designer. I would go as far as to say that it is the only practical way to handle the myriad of variables that influence the performance of a sound system in an enclosed or semi-enclosed space. Knowing the limitations of a tool allows its capabilities to be fully exploited. This chapter is devoted to enabling the sound system designer to integrate computerized room modeling into the design process.

Major principles of electro-acoustic behavior have been presented elsewhere in this text. It is requisite for the sound system designer to understand how sound radiates from loudspeakers, and how it interacts with the acoustic environment. Acoustical prediction software attempts to model this interaction, and provides a valuable tool that allows the sound system designer to investigate various loudspeaker selection and placement scenarios—the crux of sound system design. Selecting amplifiers and signal processing is trivial when compared to selecting and placing loudspeakers. In fact, the latter defines the former.

The principles presented in this chapter are universal in their application to room modeling programs. That is not to say that all room modeling programs are created equal. When you buy one of these programs, you are buying the knowledge, skill, experience, prejudices and assumptions of the developer. When terms like "perfect," "accurate," and "precise" are used to market an acoustical modeling program, you should run away.

The examples in this chapter were done using CATT-Acoustic™.

The acoustical prediction process takes the measured data of a loudspeaker, and injects it into a virtual environment. It is our desire that the virtual environment emulates the acoustical behavior of the actual physical space. Given the complexities of sound wave behavior in an enclosed space, one can clearly see the challenge, as well as the utility of a program that can even get close.

20.1 Spherical Loudspeaker Data

I'll begin the discussion of room modeling by focusing on the loudspeaker data. Garbage in, garbage out. Data that is collected improperly will produce erroneous results, no matter how sophisticated the program is that uses it.

Axial and polar measurements have long been used to characterize the performance of loudspeakers. Room modeling programs utilize spherical sound radiation data for modeling the behavior of a loudspeaker in a room. A spherical data set is measured over a full sphere of spaced positions around the loudspeaker, and is therefore able to characterize the loudspeaker's directivity, Fig. 20-1.

I have built a spherical loudspeaker measurement system from scratch that currently produces loudspeaker data files for many loudspeaker manufacturers for use in room modeling programs. The many years of invested time and money have provided some insights into what matters, and what doesn't with regard to loudspeaker data. Some of my assertions may surprise you. "More" is not necessarily "better."

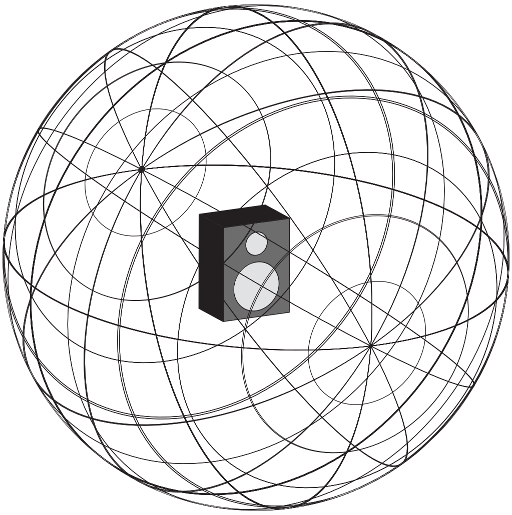


Figure 20-1. Example grid of measurement positions around a loudspeaker.

20.2 Near Field vs. Far Field

A modeling program assumes that the emerging wavefront from a source is a sphere, and that it expands spherically as the sound propagates away from the source—like inflating a balloon. This is only true in the far field of the source. The loudspeaker data must be measured in the loudspeaker's far field for the spherical spreading assumption to be correct.

The term "point source" has both theoretical and common-usage meanings in audio engineering. A literal point source is infinitely small. Since directivity is achieved by interference, and interference requires mass, a literal point source is omnidirectional by definition and would emit acoustic power that produces the same sound intensity in all directions. The spherical waves, simulated as rays or particles emerging from the point, fall off at the inverse-square law rate of level change as they propagate outward from the source, Fig. 20-2. This means that when the radius of the sphere doubles, the area that the sound passes through quadruples. Since the same sound energy is passing through a progressively larger area, the sound intensity level (L_I) and resultant sound pressure level (L_P) are predictably reduced with increasing distance.

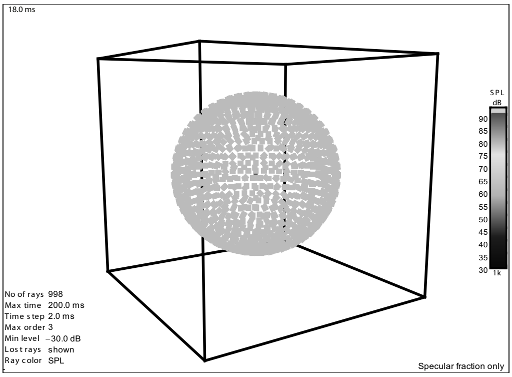


Figure 20-2. Emitted rays from a sound source in a box.

A physically realizable loudspeaker has size and mass. The sound may not radiate evenly from all of its surfaces, sides and edges, so the wavefront may not be spherical near the source. This is definitely true of multi-way devices as well as line arrays. Even though the wavefronts from these devices are not spherical when they are formed, the sound travels at the same speed in all directions. This means that the waves become increasingly spherical as the distance from the source increases, since the propagation distance, which is equal for all rays, swamps out any differences in the length of rays near the source. The distance from the source at which the waves can be considered to be spherical is the beginning of the far field. The inverse-square law applies from this distance outward. So, all loudspeakers obey the inverse-square law at remote distances, but not necessarily up close. Note that "spherical wave" does not mean that the loudspeaker is omnidirectional. While the balloon shape is spherical in the far-field, the sound from a useful loudspeaker is likely more intense in the axial direction due to the use of horns, wave guides or baffles, Fig. 20-3.

A loudspeaker has a near field where the emerging wavefront is not spherical. It has a far field where it is. There is a frequency-dependent transition between the near field and the far field. The loudspeaker's axial transfer function is distance-dependent in the near field. It is independent of distance in the far-field, except for the frequency-dependent effects of air absorption.

There are both low frequency and high frequency criteria for the beginning of the far field. To be in the far field the point of observation must be:

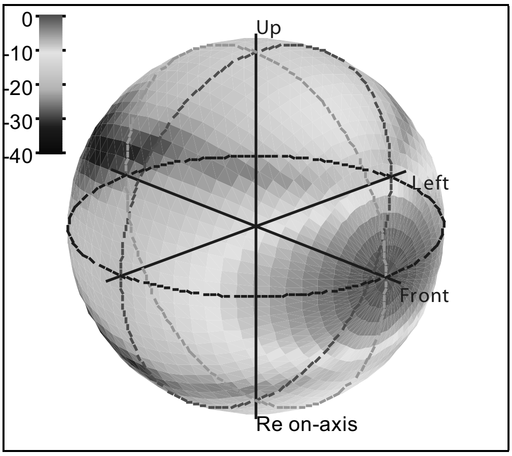


Figure 20-3. Spherical radiation with non-uniform intensity.

1. At least one wavelength from the source at the lowest frequency of interest. This satisfies the low frequency criteria.
2. At least 10 times the longest dimension of the source normal to the aiming axis. This satisfies the high frequency criteria. This assumes that the high frequency sound energy emanates from the entire surface of the device. Often it does not, and the 10 times distance criteria can be relaxed.

Using these criteria, a 1 meter (m) tall full-range loudspeaker radiating high frequency energy from its entire length would have a far-field that begins at approximately 10 m for frequencies above 30 Hz (10 m is the acoustic wavelength of 30 Hz). One can immediately see the challenge to collecting accurate spherical far field data. In practical cases this required distance can be relaxed. If uncertain, it can be determined by measurement with a pull-away test that compares the axial response of measurements made at varying distances from the source. If the axial frequency response magnitude changes with increasing distance (other than losses due to air absorption), the measurement position is in the near field.

A practical distance for measuring spherical data is approximately 8 m. Significant issues arise for lesser or greater distances. 8 m allows measurements above 43 Hz for a device up to 0.8 m (longest dimension of the device). If a shorter distance is used, the data may be inaccurate for the higher octave bands. If a greater distance is used, air absorption and thermal gradients become serious issues, especially if phase data is being measured. If a device is too large to be measured at 8 m, it can sometimes be broken down into smaller elements that are independently measured and re-assembled in software. Line arrays of discrete sources are an example, Fig. 20-4.

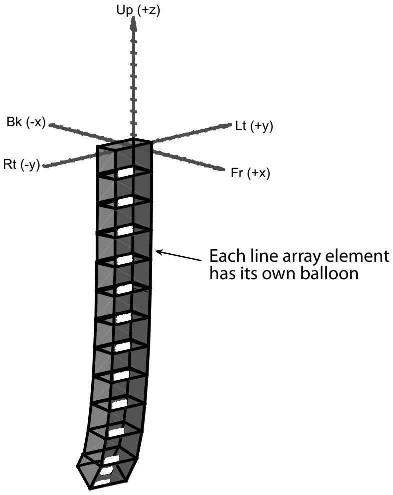


Figure 20-4. Line array of discrete sources (Nexo GEOTM). (*Courtesy CATT-A.*)

In practice, devices up to 2 m in length can often be measured at 8 m. This is because many (most) loudspeakers do not emit significant high frequency energy from their entire frontal area. The 10x distance criteria can be relaxed with an acceptable loss of high frequency accuracy. For room modeling purposes, data is only needed through the 8 kHz octave band for predicting speech intelligibility, Fig. 20-5.

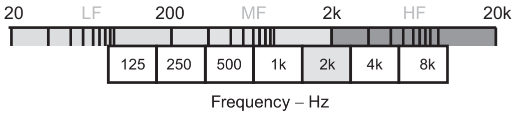


Figure 20-5. Octave bands that can be meaningfully simulated in room modeling programs.

20.3 The Measurement Process

The radiation properties of a loudspeaker must be determined by measurements made on the surface of the far field sphere previously described. The loudspeaker is placed in a free field—an environment that is free of acoustic reflections. A measurement microphone is placed on-axis and in the far field of the loudspeaker. The axial impulse response (time domain) or transfer function (frequency domain) is measured and recorded. The loudspeaker is then rotated horizontally by the desired angular resolution, typically 5°, and the measurement repeated. This continues until the microphone is 180° off-axis, Fig. 20-6. The series of 37 measurements is referred to as an "arc." The loudspeaker is returned to the axial position, rotated 5° about its aiming axis, and another arc is collected. The process continues until a sufficient number of arcs have been measured to fully characterize the spherical radiation from the loudspeaker. The exact number of arcs depends on the required number of quadrants, which in turn depends on the acoustical symmetry of the loudspeaker, Fig. 20-7.

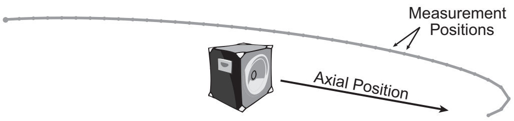


Figure 20-6. Arc of measurement positions around loudspeaker.

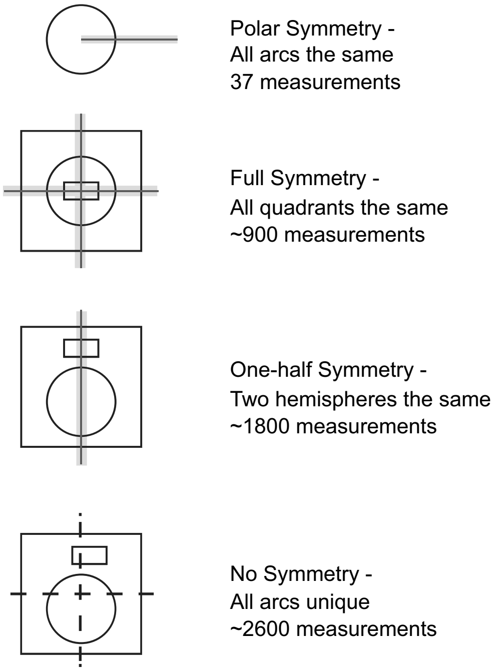


Figure 20-7. Loudspeaker symmetry and the required number of measurements.

The end result, using 5° resolution, is a set of about 2600 impulse responses. The *IR*s can be transformed to the frequency domain using the Fourier Transform to yield the transfer function, or frequency response magnitude and phase, for each measurement position. This data set is then processed into a set of loudspeaker balloon plots. There is one balloon plot for each 1/*n*-octave band. One octave resolution is generally used for room acoustics work. One-third octave may be used for loudspeaker coverage mapping and special investigations.

The balloon data is assumed by the prediction program to represent the directional behavior of the device at any distance, even though it was measured at 8 m. But, it will be inaccurate for the loudspeaker's near field. In other words, while the data balloon shape measured at 8 m may not be accurate for shorter distances in the near field, it is accurate in the far field. If the data were actually measured at one meter, and this distance is in the near field due to the loudspeaker's size, there would be a compounding error in the data as the sound propagates outward. Far field data is necessary to allow the balloon to be accurately extrapolated to remote listener positions, which are typically more abundant in large rooms than seats near the loudspeaker. We have to trade off "up close" accuracy to get "far away" accuracy. Since a main objective of the sound system design process is to achieve a positive direct-to-reverberant sound energy ratio, and since this becomes increasingly difficult with increasing distance, this trade-off is warranted.

The 1 m sensitivity of the loudspeaker (also measured in the far field and corrected to 1 m using the inverse-square law) is used by the prediction software to scale the relative balloon data to an absolute level. The room modeling program extrapolates the balloon until it intersects with an audience plane, and properly scales it to the axial sensitivity. The resultant L_P is presented as a coverage map of the audience area, Fig. 20-8. The loudspeaker data file also includes the maximum rms voltage that can be applied to the loudspeaker. I describe how this is determined in Chapter 19 *Power Ratings for Amplifiers and Loudspeakers*. The level difference between this voltage and the voltage used to measure the sensitivity is used to calculate the maximum L_P possible from the device.

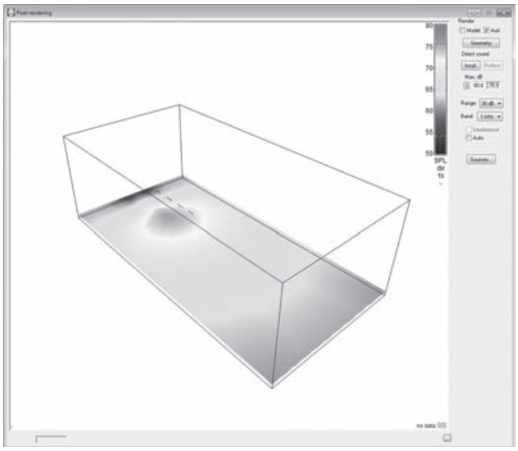


Figure 20-8. Direct field coverage map of audience area (CATT-A).

A loudspeaker that can be accurately measured and presented in this manner is commonly referred to as a point source.

20.4 Loudspeaker Arrays

Small loudspeaker arrays, such as non-steered line arrays of 1 m length or less, may be measured and modeled as point sources. Longer arrays are broken down into multiple point sources. This requires that one of the sources be measured, and then replicated in the room model. The relative arrival time of each source can be computed for any listener position within the model. This allows the complex (magnitude and phase) interaction to be calculated. Some modeling programs offer dedicated modules to facilitate array construction.

Array predictions are only estimates for a number of reasons. These include:

1. Interference of adjacent boxes. In real life, if the balloon data for a single box is measured, and a second box (no signal) is placed next to it, we have essentially created a new loudspeaker. The physical presence of the second box changes the radiation pattern of the first box, Fig. 20-9. The effect might be subtle for a pair of highly directional horns. It will be extremely significant for low directivity sources, which of course includes virtually all loudspeakers as frequency is decreased. In the array modeling program, none of the boxes know about the presence of the other boxes.
2. Consistency of array element behavior. A manufacturer typically supplies a single loudspeaker to a measurement lab for production of the data file used in room modeling programs. This loudspeaker can be measured to any practical resolution, and the trend has been to use increasingly higher angular and frequency resolution. The resolution may be as low as 10°/¹⁄₁-octave, or as high as 1°/¹⁄₂₄-octave. It is tempting to think that "more is better" and it is an easy sell to market the use of higher resolution as being more accurate. Unfortunately an inverse relationship exists here. The higher the angular and frequency resolution, the less "general" the data will be. No two loudspeakers are identical, so there is a danger of resolving a sample to the nth-degree and producing a data file that is only accurate for the sample measured, Fig. 20-10. There exists at least two possible solutions. The first is for the manufacturer to tighten their quality control standards to produce loudspeakers that have less variance from unit-to-unit. If you have ever priced a pair of matched microphones for stereo recording, you will understand the impracticality of this solution. If a manufacture must start rejecting drivers to a tight tolerance, the price goes way up. The market responds by users buying cheaper brands, so the manufacturer that is trying to "do it right" is soon out of business. The second choice is to reduce the angular and frequency resolution to something that is less sensitive to device variance. Practical resolutions are 10°/¹⁄₁-octave and 5°/¹⁄₃-octave. There are some cases where 2.5° data may be argued for a device with a very narrow beam width, such as a line array. But given the other variables that can't be controlled, this is typically unwarranted.

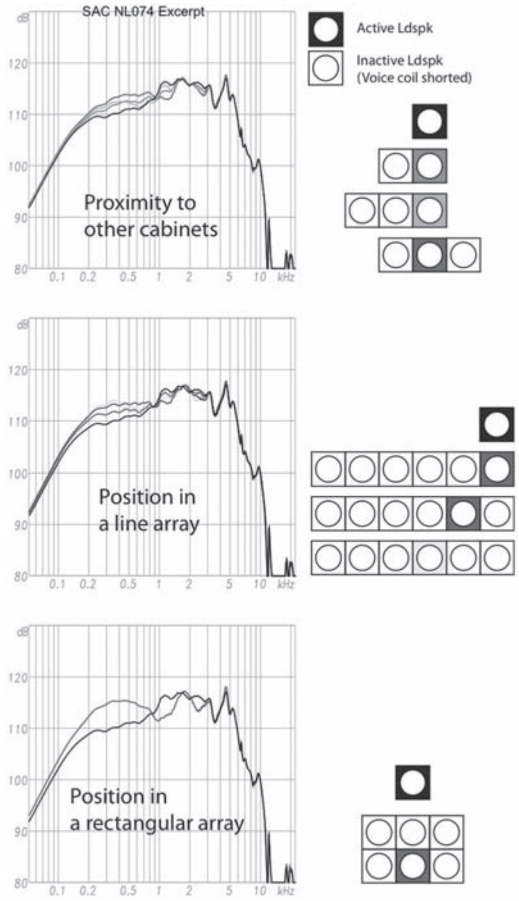


Figure 20-9. Errors caused by box interactions.

20.4.1 The Dynamic Link Library—DLL

Some array types require considerable input from the user. For example, a line array may consist of eight elements. Each must have the desired relative

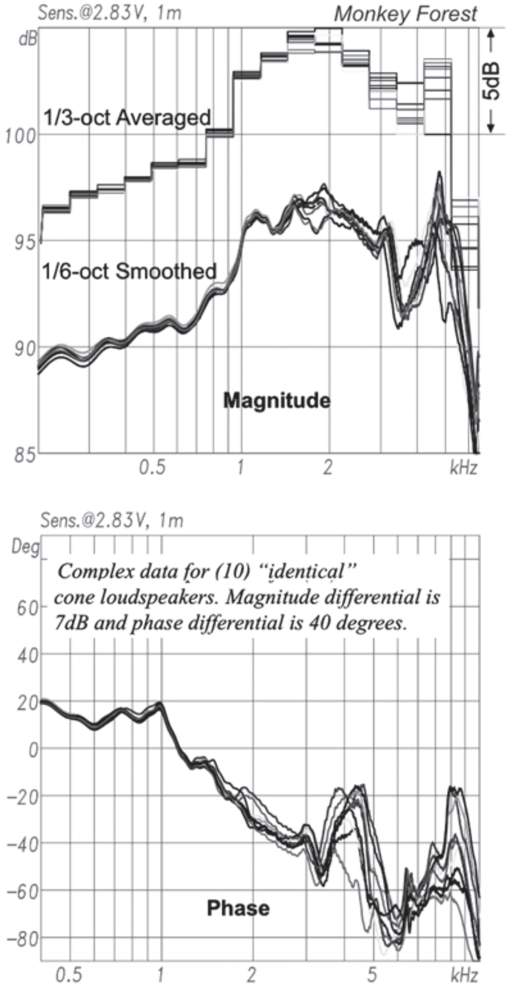
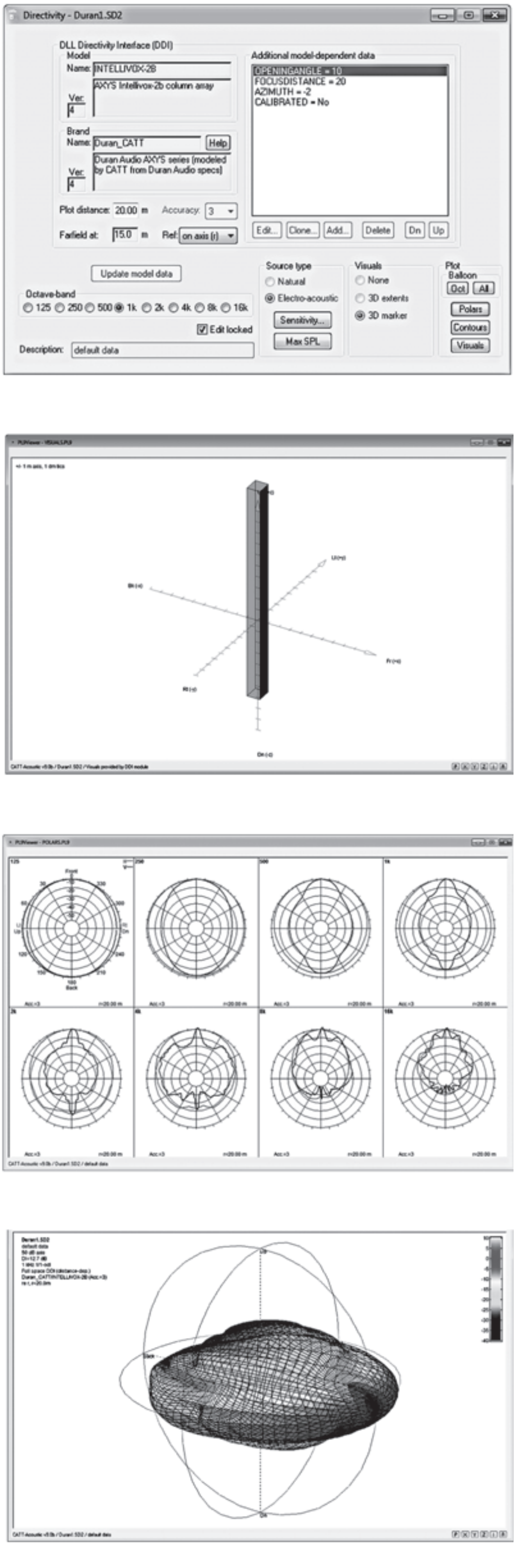


Figure 20-10. Magnitude and phase response variations of 10 "same model" loudspeakers.

position and aiming angle. Some of the elements may require delay, or custom filters. The data entry can be greatly simplified by use of a DLL. The DLL concept was first introduced in CATT-Acoustic v7.1 in 1998. Other softwares have since implemented DLLs in various forms. A DLL can allow sophisticated and complicated array configurations to be created with minimal data entry. The DLL is typically custom-coded for a manufacturer by the software developer. Fig. 20-11 shows the DLL setup for a popular line array.

20.4.2 Magnitude vs. Phase Data

A loudspeaker data balloon may consist of magnitude-only or magnitude plus phase data. Phase data can improve the prediction results for some array types, namely those that rely on complex


Directivity - Duran1.SD2

DLL Directivity Interface (DDI)

Model Name: INTELLIVOX-2B

Ver: 4

AXYS Intellivox-2b column array

Brand Name: Duran_CATT

Ver: 4

Duran Audio AXYS series (modeled by CATT from Duran Audio specs)

Plot distance: 20.00 m

Accuracy: 3

Farfield at: 15.0 m

Ref: on axis (r)

Additional model-dependent data

OPENINGANGLE = 10
FOCUSDISTANCE = 20
AZIMUTH = -2
CALIBRATED = No

Octave-band: ○ 125 ○ 250 ○ 500 ● 1k ○ 2k ○ 4k ○ 8k ○ 16k

☑ Edit locked

Description: default data

Source type: ○ Natural ● Electro-acoustic

Visuals: ○ None ○ 3D extents ● 3D marker

Figure 20-11. A DLL can allow efficient configuration of an array, and then calculate the acoustical behavior. (CATT-A).

interactions between the individual elements to form the radiation pattern, and are too large to measure as a single unit. But, phase data is difficult to measure accurately at distances that are in the far field of the source. A temperature change of a few degrees over a multi-hour measurement session can cause significant errors. If a microphone array is used to speed up the measurement time, the microphones interfere with each other in the same way that the loudspeakers in an array interfere with each other.

So, while it seems plausible to measure all loudspeakers at 1° angular resolution, and to include both magnitude and phase data, in practice the predicted response of an array is still only an estimate. If the individuals writing the software and the Standards are not involved with loudspeaker testing, or strongly influenced by those who are, there is a danger of a disconnect. Balance is brought to the discussion by considering the reason for measuring the loudspeaker in the first place—drawing board sound system design. No matter how accurate, the loudspeaker data will be used in a room modeling program that can only estimate the acoustical behavior of a space.

20.5 Direct Field Modeling

Direct field L_P and coverage predictions can be done without consideration of the room's acoustics. There is no need to build a complete, enclosed room model if all you need to know are the required mounting height and aiming angle for the loudspeaker. Modeling programs are extremely useful and accurate at showing how the spherical data balloon intersects the flat audience planes. This assures that an appropriate loudspeaker(s) has been positioned as to allow even direct field L_P over the audience area.

The direct field coverage map considers the L_P radiated from each point around the loudspeaker and the distance to the audience area. The resultant L_P is frequency-dependent, so independent maps are produced at the desired $1/n$-octave resolution. The $1/n$-octave bands can be summed and weighted to produce broadband L_P maps.

Coverage is not intuitive, and failure to model can lead to gross errors. Direct field coverage mapping should be the first step of any sound system design process, whether or not room acoustics modeling is performed, Fig. 20-12.
6+

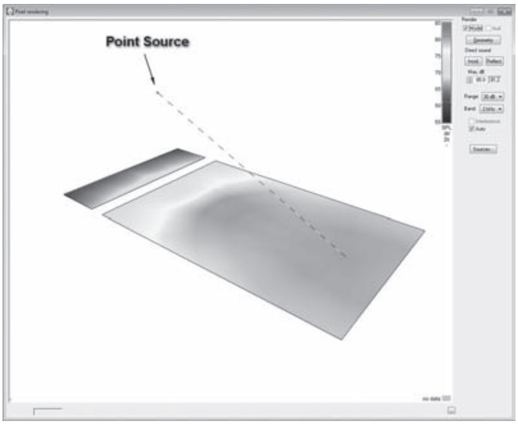


Figure 20-12. Direct field coverage map for 2 kHz octave band—audience area only.

20.6 Room Model Detail

Once the designer is satisfied that the direct field coverage is acceptable, the reflected energy from the room must be considered. A virtual wire-frame model of the room geometry is constructed for this purpose.

The amount of detail required in the room model is the subject of ongoing debate. It would seem intuitive that an accurate visual model is an accurate acoustical model. This would also allow the convenience of using an existing CAD model as provided by an architect. Unfortunately this is generally not the case. A good visual model is not necessarily a good acoustic model.

The interaction of sound with room objects is quite complex, being a combination of reflection, resonance and diffraction. Ray tracing or image-source methods can only approximate the behavior of sound, so they can never be described as "accurate" regardless of the detail used in the model. Too much detail can dramatically increase the calculation time, without increasing the accuracy of the predictions. The room model should be viewed as an acoustic "sketchpad," and is the acoustic equivalent to the architect's scale model made from foam board and paper, Fig. 20-13. It is a highly programmable reverberation processor that can be tailored to the room's geometry while considering the directivity of the source.

The accuracy of acoustic predictions tends to follow the nature of the sound from the source. The direct field emitted by the loudspeaker can be measured with very high accuracy, and its behavior in the far free field can be accurately estimated using the inverse-square law. This means that direct field measures, such as L_P and coverage can be predicted with high accuracy. Once the sound reflects, the term "accuracy" no longer applies. We are now dealing with approximations, as each ray that encounters a room boundary produces a new acoustic source that is as complex as the original source. The behavior of the sound becomes

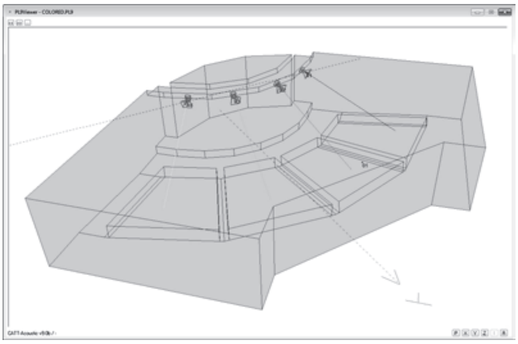
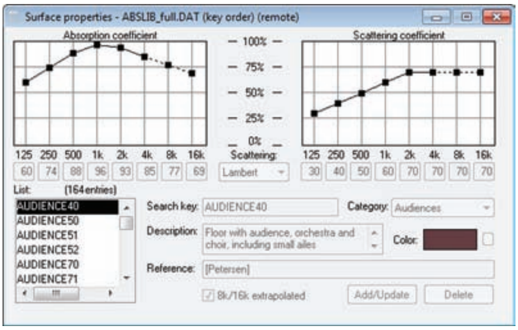


Figure 20-13. Simple 3D surface model of a fan-shaped room, showing proposed loudspeaker positions.

increasingly complex with each reflection order, eventually becoming diffuse and defying deterministic prediction completely. Fortunately, the needs of the sound system designer tend to track this accuracy progression. I can have high confidence in the direct field predictions and possibly a few reflection orders if the surfaces are large and smooth. But, the errors compound as the sound propagates and the higher order reflections and the reverberant field behaviors are only estimates.

Modeling programs usually provide specialized modules for creating the 3D surface model. Alternately the model may be created in a third party CAD program and imported. Each room boundary (or plane) is given an absorption coefficient that determines the how much the L_P of the reflection is reduced when the ray encounters a room surface. The plane can also be given a scattering coefficient, which randomizes some of the reflected energy. Scattering coefficients are invaluable for estimating the behavior of complex room surfaces. Both absorption and scattering coefficients are estimates, and their determination is one of the major challenges in the modeling process, Fig. 20-14.



Surface properties - ABSLIB_full.DAT (key order) (remote)

Absorption coefficient

125	250	500	1k	2k	4k	8k	16k
60	74	88	96	93	85	77	69

Scattering: Lambert

Scattering coefficient

125	250	500	1k	2k	4k	8k	16k
30	40	50	60	70	70	70	70

List: (164 entries)

AUDIENCE40
AUDIENCE50
AUDIENCE51
AUDIENCE52
AUDIENCE70
AUDIENCE71

Search key: AUDIENCE40

Category: Audiences

Description: Floor with audience, orchestra and choir, including small ailes

Color:

Reference: [Petersen]

☑ 8k/16k extrapolated

Figure 20-14. Absorption and scattering coefficients at ¹⁄₁-octave resolution.

20.6.1 Predicting Room Reflections

Room modeling programs primarily use two methods to predict room reflections. An image-source method is deterministic, and can be visualized by considering the room surfaces to be mirrors. If you were seated at the listener position, the image of the loudspeaker would be visible in each boundary that produces a specular reflection for your seat. Similarly, in the real room if the source were replaced with a laser-like directivity, and the images replaced with mirrors, the laser dot would end up on you. So, by using optical principles and geometry, the surfaces producing specular reflections can be identified, Fig. 20-15.

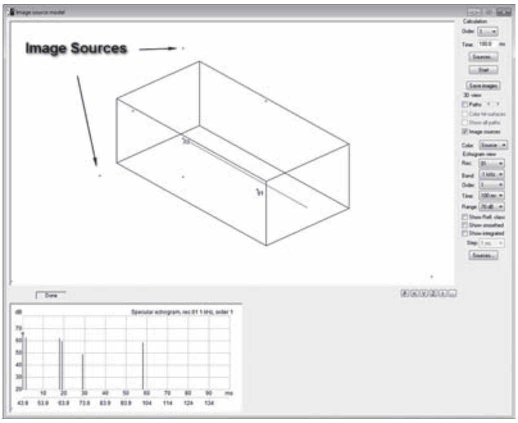
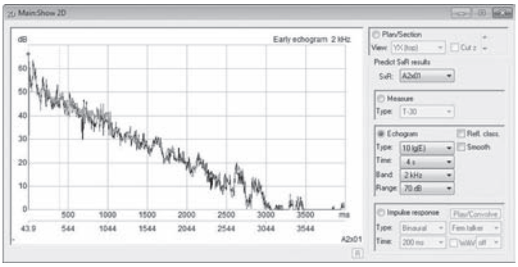


Figure 20-15. First order image sources for a single source at a single listener position. The level and arrival time for each source are shown in the echogram (CATT-A).

The computational intensity of this method increases as the reflection order is increased, while the accuracy decreases with each successive reflection. This suggests that a different approach is needed to simulate the late decay of the room. A ray tracing algorithm and its variants (e.g., cone tracing) emits thousands of virtual rays from the source, "traces" them to a user-specified reflection order, and then counts the ones that arrive at the listener position. The listener in the model is actually a "counting balloon"—a target sphere of fixed or variable radius. The exact method of predicting the reflected field differs between modeling programs, as do the results. The method used may be unique to a modeling program, and is influenced by the knowledge, skills, intuition and prejudices of the programmer, Fig. 20-16.



View: YX (top)
SxR: A2x01
Type: T-30
Type: 10 lg(E)
Time: 4 s
Band: 2 kHz
Range: 70 dB
Type: Binaural
Fem talker
Time: 200 ms
off

Figure 20-16. Full echogram for a single listener position (2 kHz octave band) (CATT-A).

20.6.2 The Objective of Room Modeling

For any given source/listener combination in an enclosed or semi-enclosed space, there are a myriad of variables that determine the room impulse response (*RIR*). The *RIR*, in turn, is the best summary that we have regarding what that seat sounds like. In a physical room, the collection of the *RIR* is the single most important task for the investigator, for within it are found the physical reasons for the sound quality at that listener position.

The objective of the design process is to synthesize an approximation of the *RIR*. There are numerous similarities and parallels between the synthesized *RIR*, which in modeling may be referred to as an echogram, and the actual measured *RIR* for the same seat in the physical room. For one to be a good modeler, one must first be a good measurer. It is through an appreciation for the sensitivities of measured data that one can grasp the difficulty of predicting the *RIR*. Only then can we avoid wasting time on the fine details and subtleties that many assume are accounted for by the mathematics used by the modeling software.

The measured *RIR* provides a reference for creating a virtual environment whose behavior emulates the actual room to the degree necessary to select and place loudspeakers. Measuring and modeling go hand-in-hand, and what is learned about one can aid one's understanding of the other.

The following sections apply to both measured and modeled acoustical data.

20.7 Room Acoustics—An Overview

The speed of sound perturbations that propagate through air is very slow compared to the speed of light. This results in a human-detectable time offset between the direct sound from the loudspeaker arriving at the listener and the reflections produced from room surfaces within the space. The art and science of room acoustics deals with these reflections.

The room is passive and produces no sound of its own. When sound is produced in the room from a source, reflections are produced by the various room surfaces. Rooms are initially analyzed in the time domain since the arrival of reflections is a function of time. For any listener position, there exists a ratio between the direct sound and the reflected "room sound." This ratio influences how well the information from the sound system is conveyed to the audience. Let's look at how the room response is evaluated.

While the "frequency response" is a much more popular way of describing the sound from a system, the frequency response is determined by the time response—the complex interaction of the direct field and multiple reflections that are unique for each listener position. Most signal processing tools (i.e. equalizers) are "time blind" and therefore fail to address the real causes of poor sound quality. Since an equalizer affects all of the sound fields heard at a listener position, it cannot alter the ratio between them, which is the root of most sound clarity and speech intelligibility problems.

20.7.1 The Room Impulse Response—RIR

A hand clap in the space will produce a series of reflections at a listener position, Fig. 20-17. Each reflection is a modified facsimile of the original event. This series can be broken down into several distinct sound fields. The hand clap results in a crude *RIR*. In formal investigations the hand clap is replaced by methods that are calibrated and consistent. It is important to understand that regardless of the method used to collect it, the *RIR* is the most fundamental acoustic test. It is the primary means of analyzing the acoustic behavior of a room, and its synthesis is the ultimate goal of the design process. The room will have the same affect on any sound coming from the loudspeaker, that it has on the impulse that was sent out. A 3D surface scale model of the room provides a virtual environment for sculpting the synthesized *RIR*.

The sound fields that the impulse produces include

1. The Direct Sound Field.
2. The Early-Reflected Sound Field.
3. The Late-Reflected Sound Field.
4. The Reverberant Sound Field.

For pure acoustics work, the impulse may be a balloon pop or starter's pistol. For sound system

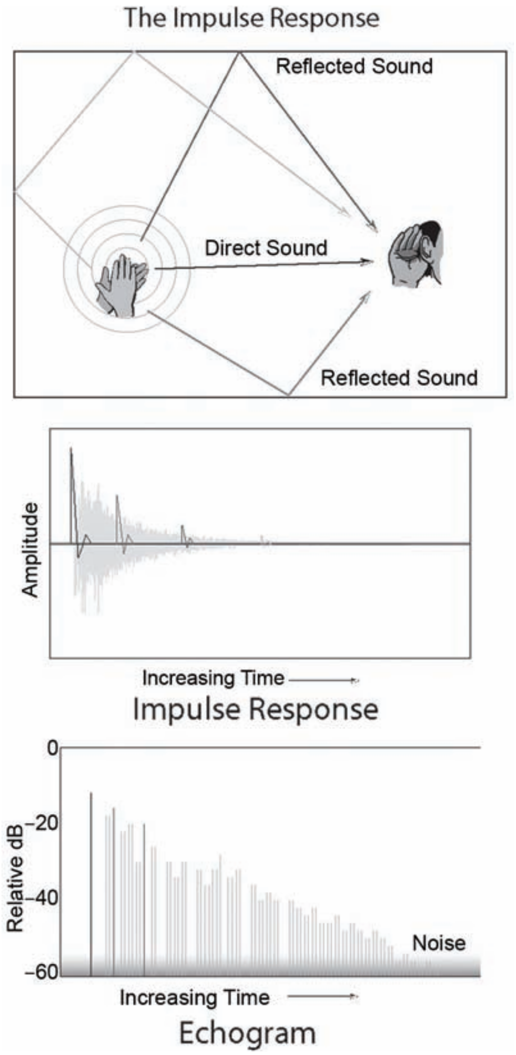

The Impulse Response
Reflected Sound
Direct Sound
Reflected Sound
Amplitude
Increasing Time
Impulse Response
0
-20
-40
-60
Relative dB
Noise
Increasing Time
Echogram

Figure 20-17. Hand clap and resultant reflections.

work the stimulus may be pseudo-random pink noise or a sine wave sweep that is played through a loudspeaker, recorded and mathematically processed to yield the impulse response by an analyzer. This technique allows the impulse response to be collected without emitting an actual impulse with it's attendant drawbacks. These include the requirement for a very quiet room as well as possible damage to the loudspeaker if the level is too high.

In the same way that a sound source and receiver position are placed in a physical room to measure the *RIR*, a virtual source and receiver are placed in the computer model to predict the *RIR*. The room model serves as a virtual measurement environment.

The *RIR* is shown in Fig. 20-18. It is a time domain plot where sound pressure (in this example) is the dependent variable and time is the independent variable. The vertical axis is linear.

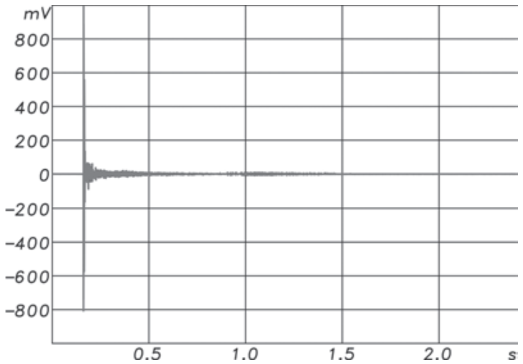


Figure 20-18. The Room Impulse Response—*RIR*.

20.7.2 Post-Processing the RIR

To aid in evaluating the *RIR*, the absolute values of the amplitudes are displayed on a vertical axis of relative dB. This is the log-squared *RIR*. All log-squared *RIR*s have some similar attributes, allowing the sound arrivals to be classified. I will represent it with a general plot, Fig. 20-19. Please pause and carefully consider the plot shown. It is the time domain representation of the relative levels of sound arrivals produced by an impulsive source placed at a unique position in the room, and collected from a different unique position.

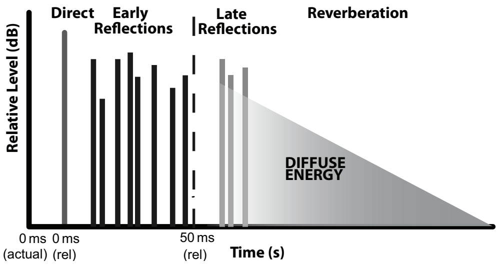

Direct
Early Reflections
Late Reflections
Reverberation
Relative Level (dB)
DIFFUSE ENERGY
0 ms (actual)
0 ms (rel)
50 ms (rel)
Time (s)

Figure 20-19. General representation of the log-squared *RIR*.

Since an infinite number of such positions exist, the investigator must select each based on what they are investigating. The answer to "Why?" determines "Where." Strategically meaningful source and receiver positions are selected for both making room measurements and for evaluating the system performance in a computer model. Both are heavily influenced by answers to three questions:

1. "Where can a loudspeaker(s) be placed?"
2. "Which model(s) will I use?"
3. "Where will the listener(s) be located?

The *RIR* is profoundly affected by these decisions, which is the whole point of using computer modeling to estimate the performance of the system at the drawing board. A "measure" or "metric" is a score used to quantify some aspect of the *RIR*. One popular basic measure is the reverberation time in seconds.

20.8 Absorption

A sound absorber terminates the sound wave at a room surface by converting part of the sound energy into heat. A rating that describes the absorptiveness of a material is the absorption coefficient or ABS. The ABS is a number between zero and one, where zero is a perfect reflector and one is equivalent to an open window from which no sound returns. The ABS is frequency-dependent and is usually specified at one-octave resolution, Fig. 20-14.

In both the actual room and the room model, the ABS is multiplied by the surface area of the boundary to produce the sabins of absorption contributed by that boundary to the space, named for Wallace Clement Sabine, considered by many to be the father of architectural acoustics. Each surface can have only one ABS, and large surfaces can be subdivided if they have multiple coverings. One English sabin is one square foot of open window. One metric sabin is one square meter of open window. The more sabins, the less reflected sound. Sabins can be added or removed to modify the sound of the room.

If all room surfaces have an ABS = 1 for all octave bands, the room would be anechoic, or without echoes. An anechoic chamber utilizes heavy ABS to simulate this condition. In practice complete absorption is not possible and one criteria for an anechoic environment is that all reflections must be –20 dB relative to the direct sound at the measurement microphone.

Absorptive materials have an acoustic impedance similar to air. Effective absorbers include soft, fuzzy materials such as fiberglass and mineral wool as well as some types of foams. Low frequency absorption may be accomplished by diaphragmatic action, a sometimes unintentional result of gypsum board or thin panel wall construction. Adding surface relief increases the surface area and yields more sabins for a given area. The ideal minimum thickness of an absorber is ¼-wavelength at the lowest frequency of interest. This makes low frequency absorption difficult to accomplish for practical reasons due to the required material thickness.

ABS coefficients are most often measured in a reverberation chamber. This accounts for the energy loss (or more correctly, the conversion to heat) of sound striking the material at random angles. The material's effect on a specular reflection may be quite different. Remember, these are estimates. Be conservative if you have to guess at an ABS coefficient.

20.8.1 Modeling with Absorption Coefficients

The absorption coefficient also plays a central role in the computer model. Here, the coefficient is a percentage. Each surface in the room model is given an absorption coefficient, usually taken from tables of values created from actual measurements of various surface coverings, often with obscure origins where the details of how they were measured are rarely given. Practical values range from 1% to 99%. A practical resolution is ⅟₁-octave. While there is a push toward higher (i.e. ⅓-octave) ABS data, "more" is not necessarily "better" for several reasons. These include:

1. No matter how complete your ABS materials library, there is always some educated guesswork involved in assigning a coefficient to a room surface. ⅓-octave resolution complicates the guessing process, requiring the designer to guess at 3 times as many values.
2. Sharp spikes in the absorptive characteristics of a material tend to support the case for higher frequency resolution, but such spikes (or dips) tend to be very sensitive to the variables that cause them. Even high resolution ABS data may not properly quantify the behavior of such a surface.
3. The predicted performance of both loudspeakers and rooms must be done at a practical resolution. It is a time-intensive task to characterize the direct field sound coverage and various performance measures at ⅟₁-octave resolution, let alone ⅓-octave.

The "appropriate resolution" discussion also applies to loudspeakers. Well-behaved loudspeakers typically do not have sharp directivity changes as function of frequency. One can usually visually interpolate the loudspeaker's performance between the ⅟₁-octave bands. Ironically, the poorest loudspeaker designs can require the highest measurement resolution to characterize, as higher angular resolution may be required to resolve their erratic response. Is is better to increase the resolution to quantify such a device, or select a device that can be characterized by a lower angular resolution? The debate goes on.

20.8.2 Sound Behavior

When the acoustic wavelengths are short relative to the surface sizes that take part in the reflection, sound is modeled as rays of light that behave geometrically. The angle of incidence will equal the angle of reflection, as with a mirror. As frequency decreases, the light model becomes increasingly inaccurate and the sound must be modeled as a wave, which is much more difficult to handle computationally. The transition of a room from "ray behavior" to "wave behavior" is a gradient with no clear single transition frequency. The "ray assumption" is why we can use straight lines with arrows to indicate the direction of sound travel. One must always be mindful of the limitations of their assumptions. It's a big jump to consider the sound as a wave rather than as a ray, which is why most acoustic prediction methods don't work well at low frequencies, Fig. 20-20.

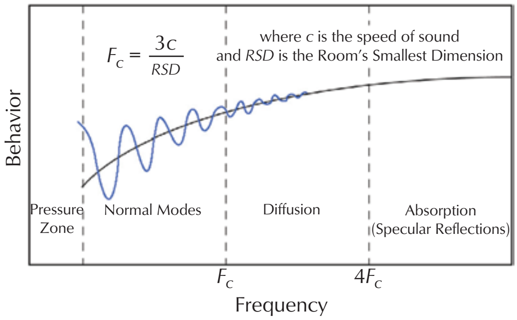
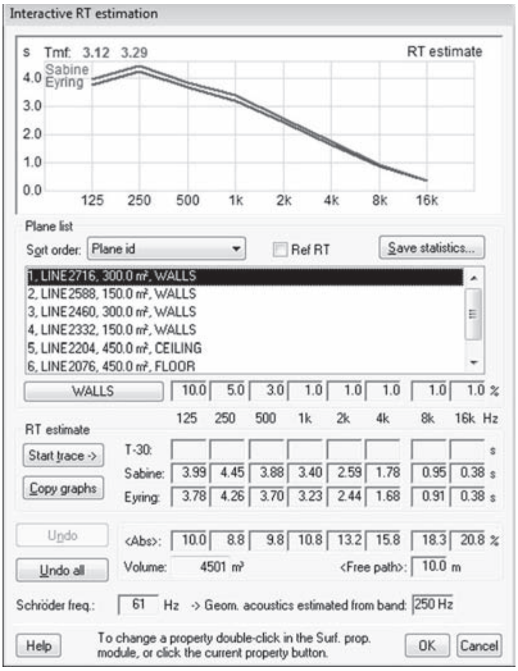

$F_c = \frac{3c}{RSD}$
where c is the speed of sound
and RSD is the Room's Smallest Dimension
Behavior
Pressure Zone
Normal Modes
Diffusion
Absorption (Specular Reflections)
F_c
$4F_c$
Frequency

Figure 20-20. Critical frequency chart, showing a simplified equation. Room modeling works best in the absorption region, and its accuracy diminishes with decreasing frequency.

The Schroeder Frequency estimates the transition region from "wave" behavior to "ray" behavior. In practical cases, the ray model for sound has validity to possibly 100 Hz, and even then only for a very large space. This means that the octave band centered at 125 Hz is the lowest band that can be considered in the room model. At the other end of the spectrum, the 8 kHz octave band extends to beyond 10 kHz. The very short wavelengths beyond the 8 kHz octave band may not predict well. They also have little if any influence on speech intelligibility.

Room modeling programs assume geometric behavior of sound, treating it as a ray or particle that is emitted from a source. For this reason, sound system performance predictions, including room acoustics predictions are best limited to the 125 Hz through 8 kHz octave bands, see Fig. 20-5.

The following discussion assumes that the room size is very large, and the individual surfaces large and smooth, relative to the wavelength of the sound, and therefore the sound behaves geometrically. This is usually the case for most of the audible spectrum for rooms sufficiently large to require a sound reinforcement system. The computer modeling programs used by sound system designers assume geometric sound behavior, and should therefore restrict acoustic calculations to the 125 Hz-octave band and above. They should also warn you regarding the useful low frequency limit, Fig. 20-21.



Tmf: 3.12 3.29

Sort order: Plane id ☐ Ref RT

1, LINE2716, 300.0 m², WALLS
2, LINE2588, 150.0 m², WALLS
3, LINE2460, 300.0 m², WALLS
4, LINE2332, 150.0 m², WALLS
5, LINE2204, 450.0 m², CEILING
6, LINE2076, 450.0 m², FLOOR

	125	250	500	1k	2k	4k	8k	16k	Hz
WALLS	10.0	5.0	3.0	1.0	1.0	1.0	1.0	1.0	%
T-30:									s
Sabine:	3.99	4.45	3.88	3.40	2.59	1.78	0.95	0.38	s
Eyring:	3.78	4.26	3.70	3.23	2.44	1.68	0.91	0.38	s
<Abs>:	10.0	8.8	9.8	10.8	13.2	15.8	18.3	20.8	%

Volume: 4501 m³ <Free path>: 10.0 m

Schröder freq.: 61 Hz -> Geom. acoustics estimated from band: 250 Hz

To change a property double-click in the Surf. prop. module, or click the current property button.

Figure 20-21. Reverberation time estimation. Note the indication of the Schroeder Frequency for this gymnasium-sized room, along with a recommended low frequency limit for acoustical predictions.

20.8.3 The Direct Sound Field

The direct field is the sound that travels straight from the source to the listener. It arrives first since it has the shortest distance to travel. The direct field is the "engineered" sound field. Loudspeaker and amplifier selection are often based on it alone, as though the system were to be used outdoors with no reflections (a free-field). The spectrum of the direct field may be equalized by modifying the drive voltage to the loudspeaker with electronic filters—a process referred to as equalization. The direct field is independent of the room, and direct field equalization can be performed off-site when certain conditions are met. The measurement of the direct field is “Step 1” of the system equalization process in real rooms, and Step 1 of the design process in the virtual room.

In the computer model, the direct field is used to model the L_P and coverage of the source over the selected room planes. The prediction is based on measured loudspeaker data as previously described.

20.8.4 The Early-Reflected Sound Field

The early-reflected field includes the reflections that arrive close enough in time to be integrated, or fused with the direct field by the ear-brain system. This integration time is frequency-dependent, ranging from a few ms at high frequencies to tens of ms at low frequencies. 35 to 50 ms is often used as a “one number” integration time.

Early reflections increase the perceived level of the sound, and may be the primary means of amplification in a lecture or recital hall. They are often called “supporting” reflections. Acousticians often use “clouds” and “shells” to provide supporting reflections to listeners or musicians. The chart in Fig. 20-22 is useful for evaluating how a reflection will be perceived based on arrival time and level.

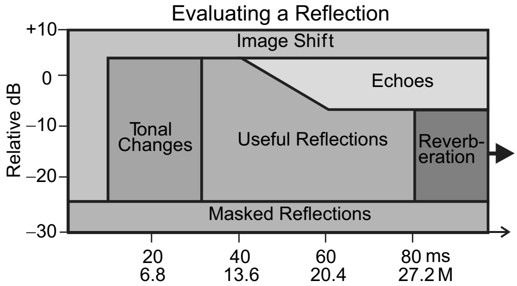


Figure 20-22. Chart for evaluating a reflection.

Early-reflections also produce tonal coloration as the reflected sound acoustically superposes with the direct sound at a listener. This can be either good or bad depending on the application. It is good sound system design practice to maintain some distance between loudspeakers and room surfaces to minimize coloration of the loudspeaker’s sound. Surface treatment can be substituted for distance, as can increasing the source’s directivity to reduce the energy striking the surface. The result is an initial time gap *ITG* between the direct sound and first reflection that can be observed on the log-squared impulse response (measured) or echogram (predicted), Fig. 20-23. The presence of an *ITG* of ten or more ms can dramatically improve the fidelity of a loudspeaker, and is considered essential in critical listening spaces such as recording studio control rooms. It is equally important in an auditorium system, where the objective is to achieve a similar response from seat-to-seat. Strong very early reflections can make this impossible (the Image Shift region of Fig. 20-22).

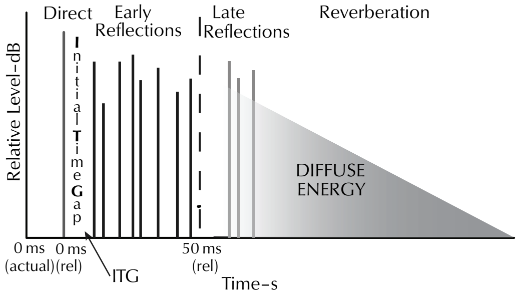


Figure 20-23. The Initial Time Gap—*ITG*.

In the room model, the early-reflected field is determined by an image-source prediction algorithm. Rather than rely on the probability of sound striking a room surface, as does ray tracing, the image-source method uses a deterministic approach. This method has computational issues for high order reflections, so most modeling programs transition from image-source methods to ray or cone tracing in order to synthesize the room echogram so that the full decay can be estimated, see Fig. 20-15.

20.8.5 Late Reflections or Echoes

Late-reflections arrive beyond the ear’s integration time. They are perceived as a blurring of the sound and in extreme cases as echoes. Hall designers use room geometry and acoustic treatment to control late reflections. Sound system designers utilize loudspeaker pattern control and placement to the same end.

In most cases, strong, late reflections that are perceived as echoes are low in order. Common offenders include rear walls, balcony faces, windows, etc. The predicted echogram can often identify the surfaces that are likely to be offending for a given listener position. As in physical rooms, the most likely problem spots—seats that receive echoes from the rear of the room—are the first rows of the audience and the stage.

20.8.6 Reverberation

The reverberation time of a room is a measure of "sound persistence." By definition, the *RT* is the time required for the sound from an interrupted source to decay by 60 dB. 60 dB was chosen for largely practical reasons, as it roughly represents the amount of decay that can be heard in a quiet space. The *RT* was more formally termed the *RT*60. This has been shortened in later standards to *T*60 to be consistent with other time domain measures. The most common measure implemented by acoustical measurement platforms is the *T*30, which is the time required for 60 dB of decay, extrapolated from the time required for 30 dB of decay. It is described in ISO 3382—"Measurement of the Reverberation Time of Rooms with Reference to Other Acoustical Parameters". The *RT* is the easiest acoustic parameter to measure, and casual investigations may only require counting or a stop watch.

Reverberation time is frequency-dependent, with one-octave being a practical resolution, see Fig. 20-21.

Reverberant energy results from the sound persisting in a space long enough to produce so many reflections that the sound becomes diffuse and random. While there is energy decay, there is no net direction of energy flow. Contrast this with a reflection which had a definitive direction of travel and that can be attributed to a specific room surface. A diffuse sound field is noise-like with regard to perception, except that it is repeatable. The two characteristics of reverberation that are of interest are the reverberation time and the reverberation level. I have previously defined reverberation time. The reverberant level is the level that the reverberant sound field builds to as the room is continuously excited by a source. A room may have a long reverberation time, but if the reverberant level is made low by careful loudspeaker selection and placement, communication may not be impaired by the reverberant field.

The reverberation time can be estimated mathematically with the Sabine equation and its variants, for rooms that meet certain criteria. These include low average absorption, uniform absorption distribution and a mixing geometric shape[1]. Few rooms meet these criteria, so the reverberation equations only provide estimates, that in some cases can have large errors. Due to its random and mixing nature, the reverberant field tends to be consistent throughout spaces where a significant reverberant field develops.

It is intuitive to think that if we just let a ray tracing or cone tracing algorithm run long enough, we will eventually get a diffuse reverberant field in the room model. This is not true. A major obstacle to predicting the reverberant "tail" in the room model is how to handle ray growth. In the real world, rays beget rays. Every encounter of a ray with a room surface essentially creates a new sound source in the room, which produces a hemisphere of rays into the room, each of which creates a new sound source when it encounters a boundary. Reverberation results when the reflections at a listener position become so dense that there is no specific direction of sound flow. Room modeling programs must estimate this ray growth, and they all do it differently. The algorithms tend to be software-specific, and sometimes more than one choice is given, depending on how the designer intends to use the data. For example, a simple, fast algorithm may be used for general *RT* estimates, but a more complex one is provided for auralization (listening). The more sophisticated algorithm may require a much longer run time.

So, the synthesis of the *RIR* transitions from deterministic number crunching for the direct field and specular reflections, to sophisticated algorithms for reverberant field. It was once popular for the reverberant field to be synthesized based on statistical reverberation equations (i.e. Sabine) and then spliced onto the end of the *RIR*. A major problem was achieving a natural transition between the specular arrivals and the dense, reverberant tail. As computers have become more powerful, so have the algorithms, and this time saving method is becoming unnecessary.

Once the computer model performs the direct field calculations, deterministic reflected field calculations, and diffuse field calculations, we have the *RIR* for the seat under investigation. The *T*30 can be determined from this data. Of all of the acoustic measures, the *T*30 is the least "seat-specific," meaning that it tends to vary less than the other measures as one moves around the space. It can serve as a general measure of sound persistence for a space, and it is meaningful to speak of working in a "3-second room." Fig. 20-21 shows a special ray tracing sub-program for quickly estimating a "global" *T*30. This is useful for achieving a general match between a measured *T*30 and the *T*30 of the model.

Reverberation is often used as a "catch-all" term for reflected sound, but it is important for sound system designers to consider the type of reflected sound as outlined here. Most rooms are semi-reverberant and all of the sound field types, plus noise, may exist at a given listener position. The investigator must determine which sound field is relevant to the problem being investigated.

20.8.7 Room Measures

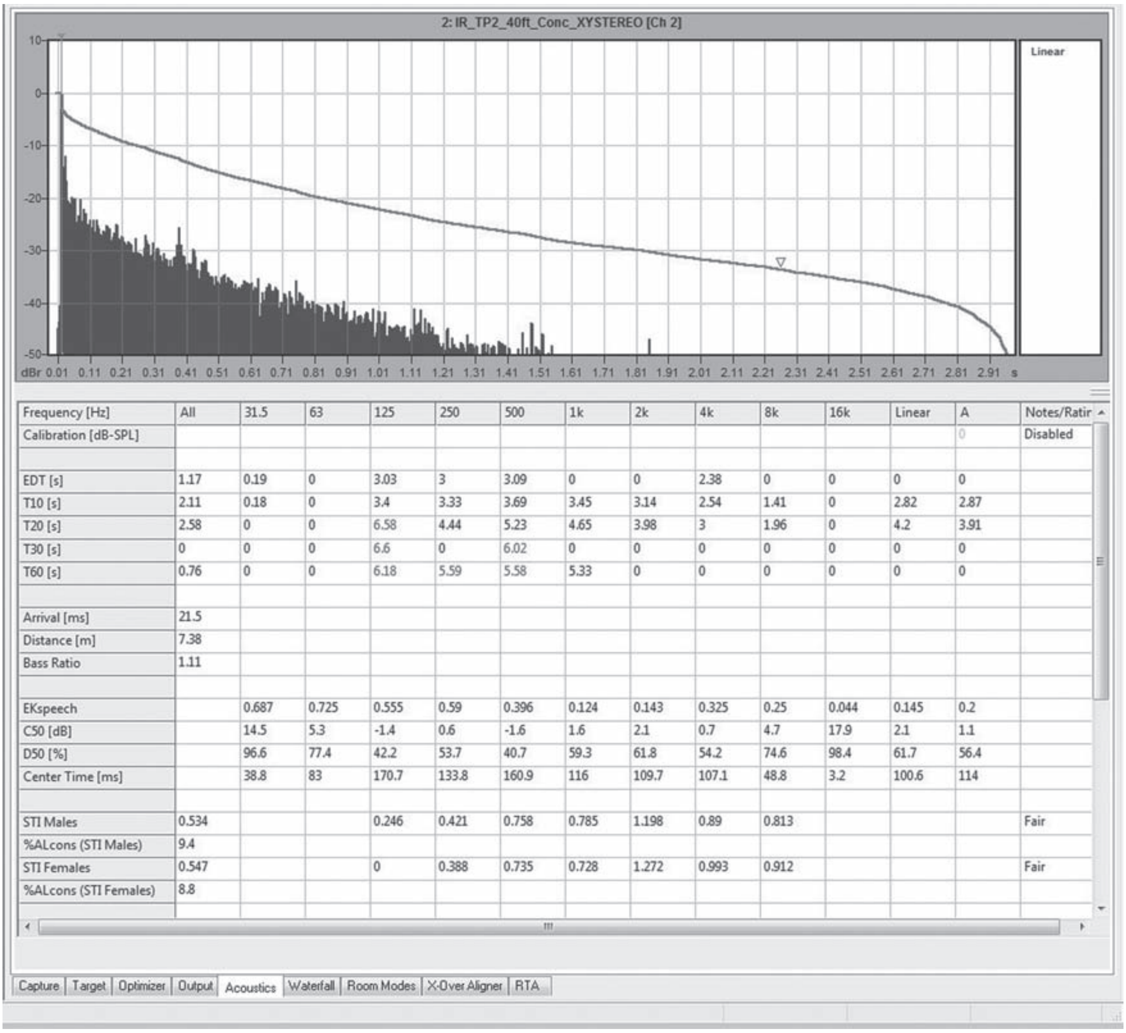
The sound arriving at a listener position can be given ratings based on the previous sound field descriptions. In actual rooms, this task is handled by an acoustic analyzer that collects the *RIR* and processes the data, Fig. 20-24.

Useful measures for sound system design include *T*30, Early-Decay Time - *EDT*, Clarity—C and one or more speech intelligibility scores.

Room acoustics is a field of study in and of itself. So is the measurement and characterization of the intelligibility of a sound system. Both are presented here as an introduction to familiarize you with their existence, and their importance in the sound system design process. Sound system practitioners must recognize the influence of room acoustics on the sound from their systems. It is possible, and even commonplace, to achieve adequate L_P and direct field coverage, only to have it swamped out by a high reverberant field level, or degraded by an echo. One objective of computer room modeling is to identify these problems at the drawing board.

20.9 Realistic Room Models

Computer room modeling programs create a virtual acoustic environment to aid in loudspeaker selection and placement. It is important to build a bridge between the physical environment and the modeled one using measured data. Skipping this step can lead to wildly inaccurate predictions that over or under


2: IR_TP2_40ft_Conc_XYSTEREO [Ch 2]
Linear
10
0
-10
-20
-30
-40
-50
dBr 0.01 0.11 0.21 0.31 0.41 0.51 0.61 0.71 0.81 0.91 1.01 1.11 1.21 1.31 1.41 1.51 1.61 1.71 1.81 1.91 2.01 2.11 2.21 2.31 2.41 2.51 2.61 2.71 2.81 2.91 s
Capture
Target
Optimizer
Output
Acoustics
Waterfall
Room Modes
X-Over Aligner
RTA

Frequency [Hz]	All	31.5	63	125	250	500	1k	2k	4k	8k	16k	Linear	A	Notes/Ratir
Calibration [dB-SPL]													0	Disabled
EDT [s]	1.17	0.19	0	3.03	3	3.09	0	0	2.38	0	0	0	0	
T10 [s]	2.11	0.18	0	3.4	3.33	3.69	3.45	3.14	2.54	1.41	0	2.82	2.87	
T20 [s]	2.58	0	0	6.58	4.44	5.23	4.65	3.98	3	1.96	0	4.2	3.91	
T30 [s]	0	0	0	6.6	0	6.02	0	0	0	0	0	0	0	
T60 [s]	0.76	0	0	6.18	5.59	5.58	5.33	0	0	0	0	0	0	
Arrival [ms]	21.5													
Distance [m]	7.38													
Bass Ratio	1.11													
EKspeech		0.687	0.725	0.555	0.59	0.396	0.124	0.143	0.325	0.25	0.044	0.145	0.2	
C50 [dB]		14.5	5.3	-1.4	0.6	-1.6	1.6	2.1	0.7	4.7	17.9	2.1	1.1	
D50 [%]		96.6	77.4	42.2	53.7	40.7	59.3	61.8	54.2	74.6	98.4	61.7	56.4	
Center Time [ms]		38.8	83	170.7	133.8	160.9	116	109.7	107.1	48.8	3.2	100.6	114	
STI Males	0.534			0.246	0.421	0.758	0.785	1.198	0.89	0.813				Fair
%ALcons (STI Males)	9.4													
STI Females	0.547			0	0.388	0.735	0.728	1.272	0.993	0.912				Fair
%ALcons (STI Females)	8.8													

Figure 20-24. Acoustics module of RoomCapture^{TM} measurement program.

estimate the clarity of the system. Following is the method that I use to qualify the room model based on a few measurements made in the physical space, assuming that the room exists.

Step 1—Field measurements

1. A well-controlled loudspeaker is placed on stage, well away from room boundaries (other than the floor). A properly measured data file for your room modeling program for this loudspeaker is mandatory, and should be provided by the manufacturer. It is not so important what loudspeaker you use as that what you use is well-defined. The loudspeaker in Fig. 20-25 is internally processed and powered, has an extremely smooth axial transfer function, and a well-defined directivity characteristic similar to a human talker. Its acoustic output is similar to that of a human talker, so a quiet room is preferred.

TalkBox
NTi

Figure 20-25. NTI Talkbox™.

General acoustics investigations are often made using a low-directivity source, Fig. 20-26. This produces maximal room excitation—exactly the opposite of what we want from a sound reinforcement loudspeaker.

2. The furthest measurement position is selected, on-axis with the loudspeaker and near the back row of the auditorium but away from room boundaries (other than the floor). The *RIR* is collected at this position.
3. The distance is halved, and the measurement repeated.
4. The distance is halved again, and the measurement repeated.

Figure 20-26. Low directivity source for room measurements. (*Courtesy Outline™.*)

I actually prefer to mark the positions first, and then begin the measurement session at the shortest distance. This prevents the measurement system from being accidentally over-driven from starting at the remote distance and moving closer to the source.

This gives me three axial positions for a well-defined source for which I can easily determine the direct field level, *RT*, clarity, speech intelligibility, etc. They are related by the inverse square law, as the direct field level should change by approximately 6 dB between these proportionally-spaced locations, Fig. 20-27. This allows the measures for in-between positions to be roughly estimated if they are needed. The *RIR* can be analyzed in a measurement program and listened to using convolution.

These measured positions are the gold standard for the performance of the room model. In a perfect world, they would be replicated exactly in the modeling environment. In the real world we will have to be content to get “close enough to dance.”

Step 2—Wire-Frame Generation

1. A wire-frame model of the room’s geometry is produced in the modeling program or other CAD environment. Each room boundary may be assigned a color and a descriptor, i.e. wall, ceiling, floor, audience plane. These will be useful later for assigning absorption and scattering coefficients. The exact procedure will vary depending on the room modeling platform.

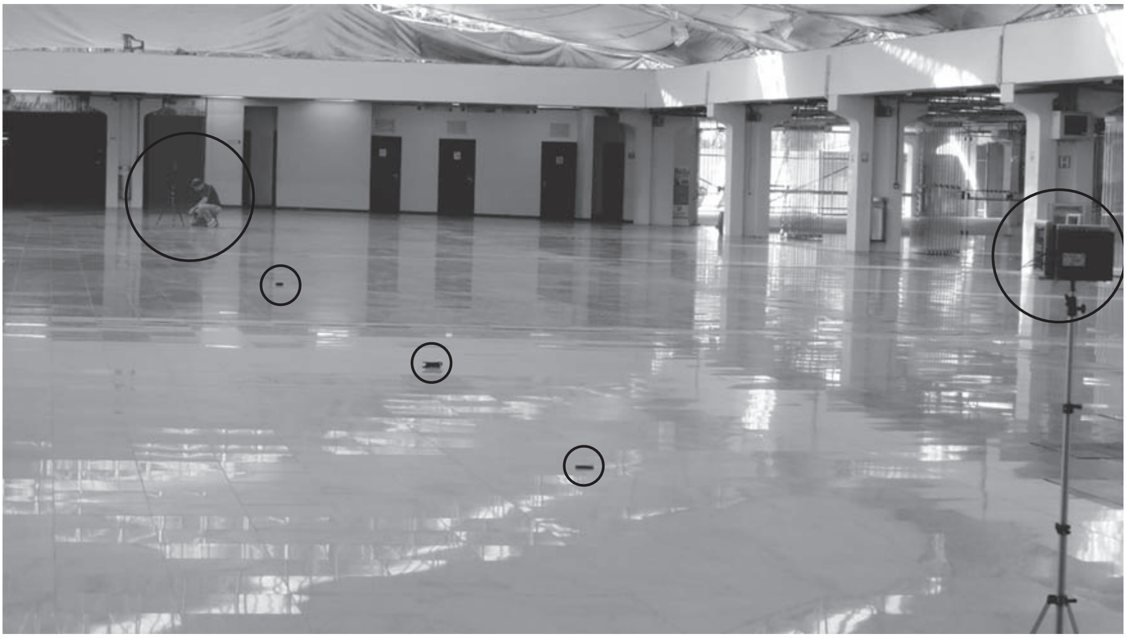


Figure 20-27. Log-spaced measurement positions.

2. In the room modeling program, I initially set the ABS for all room surfaces to 0.1, or 10%. For most spaces this underestimates the ABS, therefore overstating the *RT*. The *RT* is then calculated in the model and compared to the measured *RT* from the room. ⅟₁-octave is an appropriate resolution. Take note whether the predicted *RT* is higher or lower than the measured *RT*, as this indicates whether the average ABS must increase or decrease.
3. Refine the model by estimating the ABS coefficients for the walls, ceiling, floor etc. Since the largest room surfaces have the most influence, start with them. These values may be taken from tables of measured data, guessed at, or both. Continue this process until the ⅟₁-octave predicted values are similar to the measured values.
4. Visually identify the room surfaces with significant relief and assign scattering coefficients based on depth of the relief to the wavelength ratio for the octave band of interest, Fig. 20-28. If the relief equals half the wavelength, assign a scattering coefficient of 0.5. Increase it by 10% for each successively higher octave and halve it by 10% for each successively lower octave. This should result in a graph of scattering vs. frequency that looks like Fig. 20-14. The purpose of this step is to "de-specularize" room surfaces that will not likely produce a specular reflection due to their complex physical shape. Audiences, pews and organ pipes are examples of room surfaces that will have a high scattering coefficient.

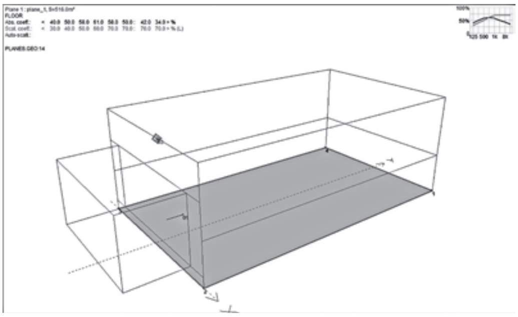


Figure 20-28. 3D surface model showing audience plane absorption and scattering coefficients at ⅟₁-octave resolution.

The use of scattering coefficients can dramatically simplify the modeling process. They allow a surface to be modeled as a flat plane, but then given a complex, frequency-dependent acoustical reflection characteristic. This allows the room surfaces that produce high level specular reflections to be more easily identified. There is no level of detail that you can use in the model to accurately quantify the acoustic behavior of a scattering surface using geometric acoustics. Attempting to do so will bog down the design process with no increase in the accuracy of the echograms.

Step 3—Correlation

We now have sufficient information to evaluate the correlation between the measured data and the room model.

1. Place a virtual source in the room model at the same position as the source in the physical room. Specify the data file for your test loudspeaker.
2. Place listener positions in the room model at the same coordinates as those used for the measured data.
3. Generate an echogram for each listener position.
4. Compare the Clarity-C50 for the measured and modeled data. Tweak the room model until the Clarity scores are reasonably close (3 dB tolerance) for measured vs. modeled, for each 1⁄1-octave band. This can take some time, and illustrates why 1⁄1-octave resolution is preferred to a higher resolution for modeling.

Step 4—Design the System

You now have a qualified room model for trying your design ideas. Substitute different loudspeaker makes and models at the positions you choose. Place additional listener seats as needed. Use coverage maps (overall) and echograms (seat-specific) to assure that the Clarity is acceptable at all listeners. Listen to your measured data using convolution, and compare it to auralizations of your predicted data. They should be similar, assuming that you use the same microphone pattern for each.

20.10 Universal Room Modeling Tips

Having been involved with the use of room modeling programs from their inception, I have learned a few things over the years that make the process more efficient and general. You do not want to turn off your brain and blindly accept what a modeling program tells you. It is providing calculated results based on your assumptions regarding the room and the assumptions of the algorithms used by the program. It if doesn't seem right, investigate further. Try to bring it back to a simplified, known condition that you can verify or correlate with measured data. Here are a few tips to aid in the process.

1. Use platform-independent programs (i.e. Sketchup™ or AutoCAD™) for producing the 3D surface model. This can allow the same model to be imported into multiple modeling programs. It also shortens the learning curve for changing modeling platforms.
2. Particle/wave duality applies. Wave methods (i.e. Finite Element Analysis—FEA) work best for low frequencies. Particle (ray) methods work best for high frequencies. There is a continuous, frequency-dependent transition between the two behaviors. Most room modeling programs use particle/ray methods for the entire spectrum. This makes accurate acoustical predictions based on geometric acoustics below the 250 Hz octave band problematic.
3. Ray tracing and image-source methods cannot fully characterize the acoustic behavior of a room surface. They provide estimates. Adding more detail to the model does not alleviate this, even though it may make the room model more visually appealing.
4. A good visual model is not necessarily a good prediction model. Too much detail can increase the calculation time without necessarily increasing the accuracy of the results.
5. The required model detail is frequency-dependent. Surfaces that are acoustically random at high frequencies may be acoustically flat at low frequencies. An example would be a large floor area covered with folding chairs. Model such a surface as a flat plane and use frequency-dependent scattering coefficients to randomize the high frequency behavior.
6. It is conceptually better to think of the room model as a highly programmable reverb unit than as an accurate acoustical model. It is an acoustical scratch pad that allows variables to be wiggled and isolated. All acoustical data (loudspeakers/room surfaces) are approximations. Accuracy and generality are mutually exclusive. The room modeling process requires educated guessing!
7. Establish a hierarchy for the design process. Here is a logical progression, along with what must be modeled to determine each.

   Step 1—Direct Field –Audience areas only.

   Step 2—Direct Field + Statistical Acoustics–Audience areas and room with correct volume and total absorption.

   Step 3—Specular Reflections—Actual room geometry. Remember that the required detail is dependent on distance from source. The rays or particles spread with distance, see Fig. 20-2. If a surface(s) is close to the loudspeaker then detail can be more important since more rays hit it. If a surface(s) is distant from the loudspeaker, detail is less important, because even with high detail few rays may strike it.

8. Break very large spaces up into smaller spaces, if they are acoustically isolated. The acoustics of the gate area of an airport typically has little influence on what is being heard in the check-in area, even though technically they may be in the same room.
9. 1⁄1-octave resolution is usually adequate to qualify a design. Higher resolutions increase the complexity of every aspect of the process, making it more cumbersome to use the model as an extension of the thought process.
10. You are building a calculator/estimator to help you quantify/qualify your ideas. You're the designer, not the computer.

20.11 Conclusions

As with making room measurements, modeling may or may not be a time-intensive task. The modeled environment allows you to do many things that are not practical or possible in real rooms. The questions raised by the modeling process may be as important as the answers that it gives. Remember to keep it simple and emphasize the majors without getting bogged down with the minors.

Philosophy is an important aspect of room modeling. One must understand what is possible to predict with reasonable accuracy and what is not. Some important attributes of the sound system designer beyond a working knowledge of electro-acoustics and room acoustics include common sense, creativity and practicality. A knowledgeable designer can likely produce a better sound system without the aid of room modeling, than a novice can with the aid of room modeling. A knowledgeable designer armed with a competently-authored room modeling program is indeed a tour de force.

Allow plenty of time to learn the modeling process before trying to use it on an actual project. Read the manual. Work through the tutorials. Start with simple, "shoebox" spaces before tackling a complex space. Over time you will develop the intuition and skill for meaningful modeling.

References

1. CATT-Acoustic User's Manual

Chapter 21

Signal Delay and Signal Synchronization

by Don Davis

21.1 Signal Delay . . . 465
    What Is Time? . . . 465
    Why Signal Delay? . . . 465
    Definitions . . . 465
    Useful Values to Keep in Mind . . . 465
21.2 Useful Signal Delay Equations . . . 466
    Temperature Effect on the Speed of Sound . . . 466
    Time . . . 466
    Travel Equations . . . 466
    Conventional Distance, Velocity, and Time Equations . . . 467
    Delay . . . 467
    A More Accurate Velocity Equation . . . 467
    Finding the Velocity of Sound with an Analyzer . . . 467
    The Henry, Fay-Hall, Haas Effect . . . 468
    A Typical Case . . . 468
    Setting Signal Delay . . . 469
    Signal Synchronization, Alignment and Convergence Defined . . . 469
21.3 Synchronization and Alignment of Arrays . . . 469
    The High-Frequency Units . . . 470
21.4 Finding Acoustic Origins of Unlike Devices . . . 470
    A Church Loudspeaker Missynchronization . . . 470
    Removal of the Reflected Energy . . . 471
    Polar Response . . . 471
21.5 Conclusion . . . 472
Bibliography . . . 474

21.1 Signal Delay

When working with sound systems, as much as one might wish to do so, we can't delay time. We can't even delay relative time. What we can and do delay is the signal. Alas, we just as often wish for a noncausal anticipating signal device as we do for a time delay. Neither is, or is likely to be, available.

21.1.1 What Is Time?

"Time flows endlessly like a river and you can't put your foot into the same river twice" so spake the ancient philosopher in bygone Greece. Clock time is based on the rotation of our planet and its base unit, the second has a specific definition, "The duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium—133 atom." *The IEEE Standard Dictionary of Electrical and Electronics Terms.*

The concept of time is nebulous and has been shown to dilate with increased velocity. No one knows the resultant velocity of the planet earth as a result of all the possible influences. Since the "big bang" or whatever other theory one prefers, we can't be sure how universal our concept of time is, but at least locally, it can be a useful parameter.

When a parameter has the dimensions in seconds (s), be sure to find out which of the many ways time is being looked at by the measurement. The words we have discussed here are quite often misused or distorted in their meaning. You don't have to wear four wrist watches (for the four U.S. time zones) and a pocket watch set to GMT plus a portable WWV receiver to be timely. Many of us rely on our Heath Master time clock synchronized to WWV and settle for the time to the nearest millisecond. Others write Greenwich and ask for the exact GMT as "they don't trust the time given out by those little southern radio stations."

21.1.2 Why Signal Delay?

If the acoustic signals had the same velocity of propagation, *c*, as the electrical signals, there would be no need for signal delay devices. The difference in velocity, unfortunately, is vast. Electromagnetic waves in free space travel at approximately 982,080,000 ft/s. Such a wave requires only 101.8 ns (nanoseconds) to cross 100 ft.

21.1.3 Definitions

0.001 s = 1 ms (millisecond)
0.000001 s = 1 µs (microsecond)
0.000000001 s = 1 ns (nanosecond)
0.000000000001 s = 1 ps (picosecond)

Acoustic waves traveling at 1130 ft/s require 88.5 ms to travel the same distance. This is a difference of 869,000 to 1.

Signal delay devices typically offer resolutions of 1.0 ms steps and precision delays of 10 µs steps.

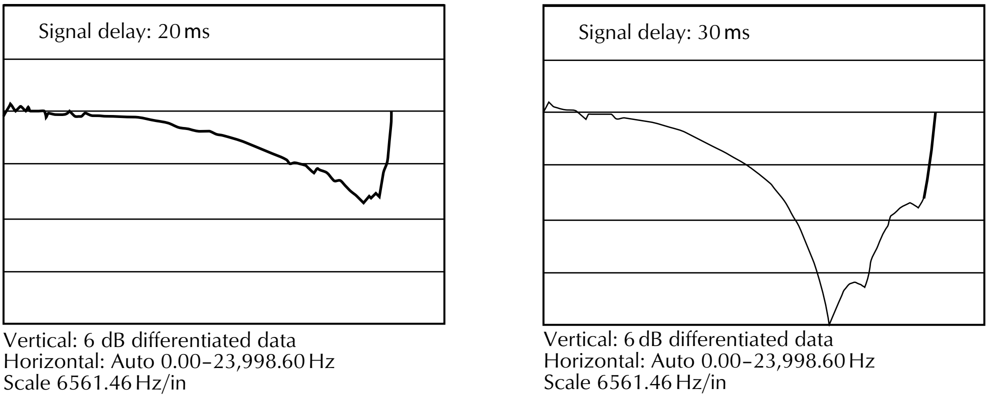
21.1.4 Useful Values to Keep in Mind

1. 1.13 ft/ms.
2. 13.56 in/ms.
3. 0.01356 in/µs.
4. 0.1356 in/10 µs, approximately ⅛ in.

Our sense of hearing can detect the results of as little as 15–20 µs difference in path length, Fig. 21-1. Very short misalignment of devices causes high frequency notches. Note that the measurements in Fig. 21-1 are differenced measurements, meaning that it was first normalized "flat" by the analyzer so that only the differenced signal was seen. The scale vertically is 6 dB/div. The horizontal scale is 6561 Hz/in (linear). Delays from as small as 100 µs to those greater than 50 ms can actually interfere with speech intelligibility. At the turn of the last century, rubber tubes of unequal length were inserted into the ears of listeners and spoken through to determine sensitivity to sound delay paths as short as 1.0 µs.

One test brought to our attention by Carolyn "Puddie" Rodgers, as well conducted and significant, was that of Hebrank, Wright, and Wilson at Duke University. "They used noise added to delay of itself and presented monaurally as their test stimuli…. They found that on A/B tests, differences in delays on the order of 7 µs could be detected." That translates into $d = ct = 1130\ (0.000007)\ (12) = 0.095$ in of acoustic path difference if that form of measurement were to be used. In far more casual experimentation, we found that large groups could easily hear the difference on speech signals of 20 µs and 30 µs.

It is common practice in sound system design to use a single point array in the front of an auditorium to cover the main audience area and overhead loudspeakers in the ceiling of under-balcony areas that are shadowed from the main array. The resulting signal delay, due to the length of the main array's longer path to the under balcony area and the very



A. 20 ms signal delay is more audible on music than on speech. B. 30 ms signal delay is 0.41 inches out of alignment.

Figure 21-1. Our hearing can detect 20 μs difference in path length.

short path from the overhead speakers, requires that the overhead loudspeakers be signal delayed to compensate for the difference in path length.

Today digital delay devices dominate the market. The delay through analog-digital and digital-to-analog converters is finite. Additionally, some digital configurations in filters can provide significant unsuspected delays. In real time sound reinforcement systems, the internal delays of all components should be known prior to inclusion. On one occasion in the author's experience, a 30 ms delay was hidden in a digital crossover network in a packaged loudspeaker system. This loudspeaker, in order to be used in the sound system, had to be placed 34 ft in front of other loudspeakers associated in the system.

21.2 Useful Signal Delay Equations

21.2.1 Temperature Effect on the Speed of Sound

$$c = 49\sqrt{459.4 + °\mathrm{F}} \tag{21-1}$$

where,
c is the velocity of sound in ft/s,
°F is the Fahrenheit temperature.

Example

$$49\sqrt{459.4 + 72.42} = 1130\ \mathrm{ft/s}$$

$$°\mathrm{F} = \frac{c^2}{49^2} - 459.4 \tag{21-2}$$

Example

$$\frac{1130^2}{49^2} - 459.4 = 72.42°$$

21.2.2 Time

$$\frac{x\ \mathrm{ms}}{\mathrm{ft}} = \frac{1}{c} \times \frac{1000\ \mathrm{ms}}{\mathrm{s}} \tag{21-3}$$

where,
c is the velocity in ft/s,
Time (in ms) = *Distance* (x ms/ft).

Example

$$\frac{1}{1130\ \mathrm{ft/s}} \times \frac{1000\ \mathrm{ms}}{1\ \mathrm{s}} = \frac{0.885\ \mathrm{ms}}{\mathrm{ft}}$$

100 ft × 0.885 ms/ft = 88.5 ms

21.2.3 Travel Equations

$$\frac{x\ \mathrm{ft}}{\mathrm{ms}} = c \times \frac{1}{1000\ \mathrm{ms}} \tag{21-4}$$

where,
c is the velocity of sound in air in ft/s.

$$\mathit{Distance} = \mathit{Time}\ (\text{in ms}) \times \frac{x\ \mathrm{ft}}{1\ \mathrm{ms}} \tag{21-5}$$

Examples

$$\frac{1130\text{ ft}}{\text{s}} \times \frac{1\text{ s}}{1000\text{ ms}} = 1.13\text{ ft/ms}$$

$$88.5\,\text{ms} \times 1.13\,\text{ft/ms} = 100\,\text{ft}$$

21.2.4 Conventional Distance, Velocity, and Time Equations

$$T = \frac{d}{c} \tag{21-6}$$

$$d = cT \tag{21-7}$$

$$c = \frac{d}{T} \tag{21-8}$$

where,
T is the time,
c is the speed of sound,
d is the distance.

Examples

$$\frac{1000\text{ ms}}{1\text{ s}} \times \frac{0.016\text{ s}}{1} = 16\text{ ms}$$

$$186{,}000\text{ mi/s} \times 0.016\text{ s} = 3000\text{ mi}$$

$$\frac{3000\text{ mi}}{0.016\text{ s}} = 186{,}000\text{ mi/s}$$

21.2.5 Delay

$$Delay\text{ (in ms)} = d\left(\frac{1}{c} \times \frac{1000\text{ ms}}{\text{s}}\right) \tag{21-9}$$

where,
d is distance,
c is the speed of sound in air.

Example

$$T_{ms} = 100\,\cancel{\text{ft}} \times \frac{1\,\cancel{\text{s}}}{1130\,\cancel{\text{ft}}} \times \frac{1000\text{ ms}}{1\,\cancel{\text{s}}}$$
$$= 88.5\,\text{ms}$$

21.2.6 A More Accurate Velocity Equation

The velocity of sound c in air for normal temperatures at sea level is dependent upon the density ρ of the air. (The temperature of the air has a major influence on its density.) The barometric pressure exerts a lesser effect under normal circumstances, the ratio of specific heats γ for air = 1.402 and the equillibrium gas pressure $P_s = 1.013 \times 10^5\,\text{N/m}^2$ (the atmospheric pressure). The velocity c of sound can then be expressed as:

$$c = \sqrt{\frac{\gamma P_s}{\rho}} \tag{21-10}$$

The variable here is the density ρ and density can be found accurately in specific cases by:

$$\rho^* = \frac{0.00129H}{1 + 0.00367(\text{K})(76)}10^3 \tag{21-11}$$

*In kilograms per cubic meter

where,
H is the barometric pressure in centimeters of mercury,
K is the temperature in Kelvins.

21.2.7 Finding the Velocity of Sound with an Analyzer

Place two identical loudspeakers exactly 1 ft apart, having first observed the 6 dB add on the ETC for the initial alignment. Two loudspeakers that are equal amplitude and equal phase, i.e., identical distances, from a measuring microphone will add +6 dB whereas equal amplitude but random phase only adds +3 dB. The +6 dB add insures the exact overlap of the two loudspeaker patterns. Because they are identical, we do know that they are exactly 1 ft apart even though we do not know exactly where their acoustic origins are. See Fig. 21-2 for an explanation of loudspeaker separation and how comb filters are generated.

Now measure the frequency interval between the "nulls" of the comb filter. This frequency interval (the lower frequency null subtracted from the higher frequency null) is the velocity of sound in ft/s. If SI dimensions are desired, separate the two loudspeakers by one meter.

This is true because the null frequency interval (NFI) is equal to the velocity (c) of the medium divided by the distance (d).

$$NFI = \frac{c}{d} \tag{21-12}$$

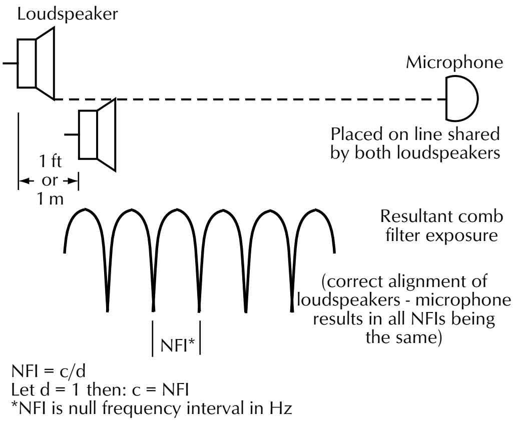

Loudspeaker
Microphone
Placed on line shared by both loudspeakers
1 ft or 1 m
Resultant comb filter exposure
(correct alignment of loudspeakers - microphone results in all NFIs being the same)
NFI*
NFI = c/d
Let d = 1 then: c = NFI
*NFI is null frequency interval in Hz

Figure 21-2. Measuring the velocity of sound using the Gold Line TEF analyzer.

Thus, if d is made unity (equal to 1.0) then:

$$c = NFI$$

The beauty of this technique is that any set of arbitrary marks on the two loudspeakers, so long as they are in identical locations on the loudspeakers, can serve as the measuring points for the determination of "d." The frequency calibration of the analyzer is not dependent upon the velocity of the media but upon the accuracy and stability of the internal clock.

It is important to align the microphone exactly on a line equal angle from the two loudspeakers, both vertically and horizontally (remember, one loudspeaker is to the rear of the other). The analyzer is "tuned" for the maximum depth of notch (i.e., halfway between the two loudspeakers). Every effort should be made to achieve a high frequency resolution, f_R, at the sacrifice of some slight reflective interference. When the *NFI* is exactly the same between the null frequencies, then you are reasonably assured that the geometric alignment between the loudspeakers and the measuring microphone is correct.

21.2.8 The Henry, Fay-Hall, Haas Effect

The Henry, Fay-Hall, Haas effect phenomenon has a lengthy history starting with Joseph Henry's remarkable experimentation at the Smithsonian in the 1840s. Henry used a child's "clicker" to listen to reflections from a large brick wall to determine the zone of inaudibility of the reflected energy.

A listener equidistant from two loudspeakers with identical levels at his or her ears will develop a phantom halfway between the two, see Fig. 21-3. However, if one of the loudspeakers has a signal to it delayed by 20 ms, then all the sound appears to come from the loudspeaker with the undelayed signal. In fact, the delayed-signal loudspeaker will have to be raised 10 dB to again seem equal to the listener. This effect occurs in anechoic chambers and over headphones from about 1 ms up to about 30+ ms. In real spaces with real loudspeakers, 20–25 ms works best.

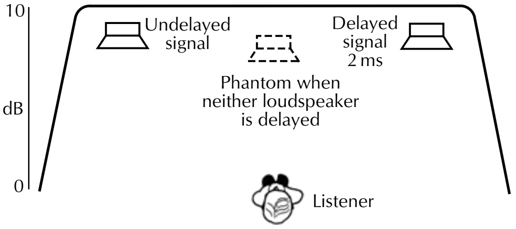


Figure 21-3. The Henry, Fay-Hall, effect.

21.2.9 A Typical Case

A typical case would be the single source system discussed earlier with the overhead under balcony loudspeaker covering the shaded areas the front loudspeaker cannot see into directly. The distance, D_2 (1), from the loudspeaker in front to the listener in the under balcony area is 100 ft. The distance, D_2 (2), from the overhead ceiling loudspeaker to the same listener is 10 ft. The needed delay is

$$\begin{aligned} T_{ms} &= [D_2(1) - D_2(2)] \times 0.885 \\ &= [100 - 10] \times 0.885 \\ &= 79.7\text{ ms} \end{aligned}$$

Best practice dictates not stopping there but adding from 5–20 ms extra to obtain the Haas effect, keeping the illusion that all the sound is from the front loudspeaker.

With single source systems the Haas effect can be obtained naturally by placement of the single source loudspeaker above and behind the talker so that D_2 exceeds D_0 by 20 ms. For a 100 ft D_0 (88.5 + 20) × 1.13 ft = 122.6 ft, see Fig. 21-4. The Doak and Bolt criteria provide a guide to determine when a delay is detrimental, see Fig. 21-5.

Where loudspeakers have to be mounted above, but well forward of the talker's position, signal delay can be used to localize the source as the talker. Listeners to such systems sometimes say the system is doing nothing and they are hearing only the talker until they are given a demonstration of the system "on" followed by the system "off" and discover that the loudspeaker was indeed 10 dB higher than the talker.

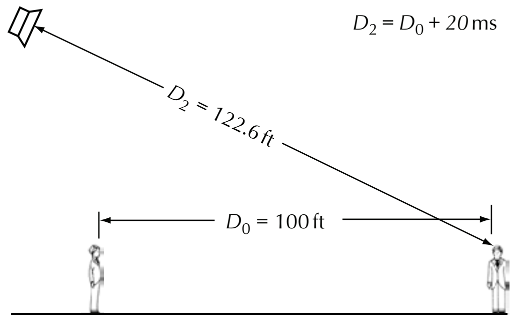
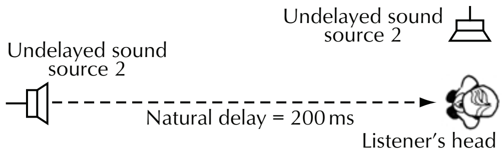
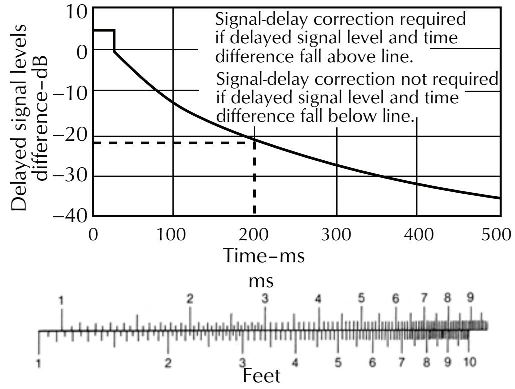


Figure 21-4. Using natural signal delay.



A. Relative delay at listener.



B. Signal delay versus relative levels.

Figure 21-5. Doak and Bolt delay-versus-level criteria.

21.2.10 Setting Signal Delay

A clever but accurate way to set signal delays is to use "clicks" at the performer's microphone with the listener wearing a headphone on one ear standing out in the audience area where the delay is to be applied. The delay is adjusted until the electrical signal, in the head phone, synchronizes with the acoustic signal through the air. The delay should be adjusted past synchronization to missynchronization from too much rather than too little delay in the delay device. Halfway between these two points on the delay device should be optimum. (Credit Rick Clarke of London.)

21.2.11 Signal Synchronization, Alignment and Convergence Defined

Signal alignment is usually a physical adjustment. It can be the attempt to physically adjust a L. F. device to a H. F. device in the crossover region. This can rarely be accomplished without a high resolution analyzer. Signal alignment can also mean two like devices, i.e., two H. F. horns, where arbitrary markings on one can be aligned with identical markings on the other.

Synchronization is usually used when electronic means are employed, usually a digital delay device, to synchronize two devices that must be physically separated but must be acoustically brought into identical arrival times at the listener's ears. Another use of synchronization is the adjustment of underbalcony loudspeaker delay to match the arrival of the single source front array. Usually such synchronization is not exact but includes the additional delay for the Haas effect to occur.

Where the pattern converges is called the acoustic center. Where the sound emits from in relative time is called the acoustic origin. Ideally we would like to have both the origins and the centers aligned whenever possible. Sometimes physical alignment of one requires that we electronically synchronize the other, i.e., long-throw horns with short-throw horns in the area where the –6 dB overlaps between their polar responses occur.

A "fly in the ointment" in either method can be convergence, where the apparent emergence of the spreading wave is different for the horizontal and vertical planes. Such a condition is called acoustic astigmatism. When such devices (they are usually high *Q* devices) are used singly it is acceptable, but attempts to array them can cause serious compromises.

Knowledge of these terms is necessary in evaluating the compatibility of devices in an overall array.

21.3 Synchronization and Alignment of Arrays

Since we cannot delay time or otherwise manipulate it, we can use audio devices to delay one signal relative to another by various methods of storage and retrieval of the signal or by adjusting the position of one source relative to another as measured at some other relative time, neither of which is dependent upon absolute time for its operation.

In signal synchronization a millisecond is a long time unit. A significant parameter to remember is 74 µs = in (73.75 to be tediously exact). The dimensions from one inch to about three feet are audible on-axis but the real danger is that they change the polar response of two devices covering the same frequency range and audience area. The lobes in many cases go straight to the "hot" microphone and are higher in level than the "on-axis" lobe.

Critics of synchronization often say that full synchronization can only be achieved at a single point in space. What's overlooked is that signal synchronization, when performed at the overlaps between devices, results in the patterns retaining their directional integrity whereas just inches of missynchronization cause the polar responses of both devices to be corrupted.

The synchronization and alignment of arrays is no longer an open question, but rather a pressing necessity. While alignment at the crossover frequency between the low-frequency devices and the high-frequency devices is indeed desirable, it is not the highest priority. The signal synchronization that is absolutely critical is the synchronization between identical or similar devices that share the same frequency range while at the same time sharing a portion of the same coverage area.

In very high quality, powerful sound systems, the frequency spectrum from 500–5,000 Hz must be carefully tailored by equalization, signal alignment, signal synchronization and coverage. In terms of coverage and signal alignment or synchronization, it is highly desirable to have devices whose characteristics are not dramatically frequency dependent over the 500–5,000 Hz range. When multiple devices must be employed in order to achieve coverage, their overlap zones must be brought into signal synchronization. Some devices won't allow signal synchronization because their acoustic centers and their acoustic origins are not compatible.

Two or more devices covering the same frequency range at the same levels but with one signal delayed relative to the other is the most serious case. It can be repeatedly demonstrated that 30 µs (thirty-one- millionths of a second) is clearly audible on speech signals over a single loudspeaker but with two signals only 30 µs apart, see Fig. 21-1. The identical effect occurs if we use two loudspeakers with one 0.41 in behind the other, i.e., 30 µs.

21.3.1 The High-Frequency Units

We take the same approach with the high-frequency devices. In this case, however, we must take great care to ascertain acoustic origins and align them because at these shorter wavelengths misalignment of acoustic origins can cause changes in the amplitude response of the system. (See Chapter 18 *Loudspeakers and Loudspeaker Arrays* for a discussion of acoustic origins, centers, etc.) If misalignment of acoustic origins simply cannot be avoided in a given case, then we must be sure it is not a small misalignment or one that causes a null frequency interval (*NFI*) to fall in the array's crossover region.

The advent of digital signal delay devices that allow signal-delay adjustments in 10 µs steps has provided array designers with enviable freedom on the geometry of the array while preserving the synchronization criteria.

The ETC display shown in Fig. 21-6A is of two measurements overlaid on the screen. The first curve shows the energy arrival in time for two like devices that are 0.170 ms (2.3 in) apart. The second curve shows the result of dialing in 0.170 ms delay to the loudspeaker that originally arrived first. Note particularly the 6 dB increase that occurs for coherent addition of the two signals once they are aligned.

The energy frequency curve, EFC, display, Fig. 21-6B, is again two measurements overlaid on the screen. The first curve shows the comb filtering produced by two like devices being 0.170 ms out of synchronization. The second curve shows the result of the 0.170 ms delay being used.

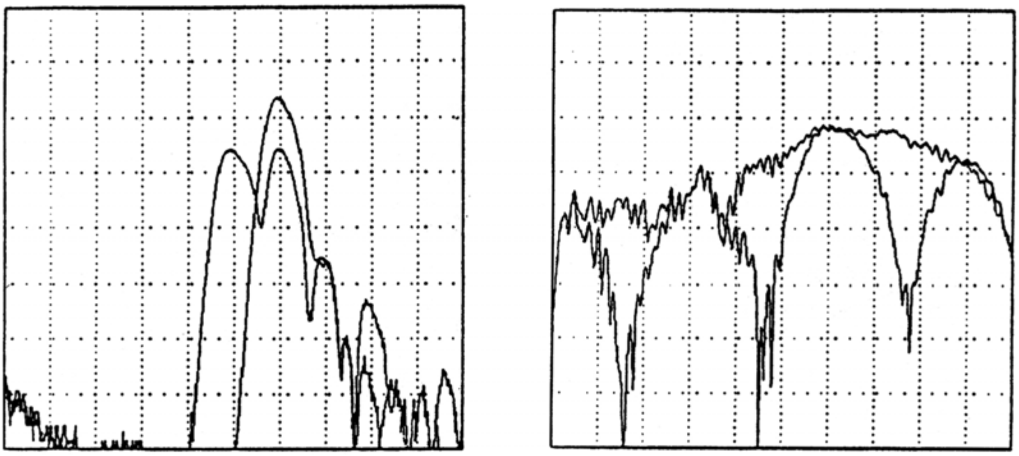
We frequently encounter a school of thought that feels that all one really has to do is get coverage. *N* factors, signal synchronization, *Q* selection, all give place to coverage. We really do not know what set of values dominated in this design of an array. Since all horns are identical, we might guess that coverage was the main consideration. The measurements tell the story, see Fig. 21-7.

21.4 Finding Acoustic Origins of Unlike Devices

When both the directional control device, DCD, and the transducer of two high-frequency units are identical, even though we do not know where their acoustic origins actually are, we can know that they are in the same place for both. This means that any arbitrary mark made in exactly the same place on both units allows accurate alignment of the two devices to each other. Whenever either the transducers or the DCDs differ from each other, then this simple method no longer is applicable. It is worthwhile to repeat that the alignment of acoustic origins is absolutely necessary when the devices being aligned cover the same frequency range. Some devices won't allow such alignment due to acoustic astigmatism, i.e., their acoustic centers and their acoustic origins are not compatible.

21.4.1 A Church Loudspeaker Missynchronization

An excellent illustration of what can occur when seemingly minor missynchronization is present is



A. ETC of two loudspeakers 2.3 inches apart—note the 6 dB gain when in alignment.

B. EFC of two loudspeakers where the bottom trace shows the comb filters from misalignment and the top response is in alignment.

Figure 21-6. Two loudspeakers misaligned.

shown in the case of a Catholic church with an RT_{60} of 2.4 s at 2 kHz.

A loudspeaker array consisting of a 15 inch low-frequency unit, a ten-cell short-throw horn, and a two-cell long-throw horn was installed in this church. In the area about mid-church where the patterns of the two horns overlapped, there was a complaint of poorer quality than seats either further forward or further back. ETC measurements revealed a 30 µs missynchronization. The EFC showed severe comb filtering in the acoustic response. The measurement of $\%Al_{CONS}$ was 10.5%.

A precision signal-delay unit was then connected to the array just ahead of the power amplifiers. The device arriving first (the near-throw horn) was connected to a delay output on the precision signal delay and was delayed by exactly 300 µs, approximately 4 inches, the far-throw device was connected to the reference output of the precision signal delay. Fig. 21-8 shows the before and after ETCs. Fig. 21-9 shows the before and after EFCs. Note the tremendous increase in L_D at the measurement point. The EFC reveals the corrected acoustic response; the $\%Al_{CONS}$ is now 6.9%.

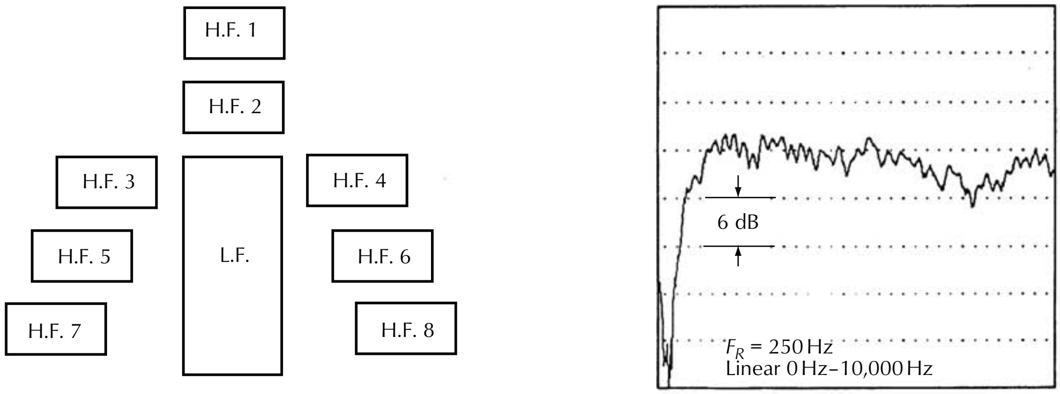
21.4.2 Removal of the Reflected Energy

The removal of the undesired polar pattern, due to missynchronization, striking the offending walls was startling. I was standing in the overlap area of the two devices and hearing sound from the surrounding walls when the correction was inserted. It was as if the walls had fallen away. Subjectively the most startling effect was the almost total removal of audible reflected energy from the room when the synchronization correction was turned on. Prior to switching to synchronization, sound from the rear wall was nearly as loud as from the array, and substantial side wall reflections were apparent.

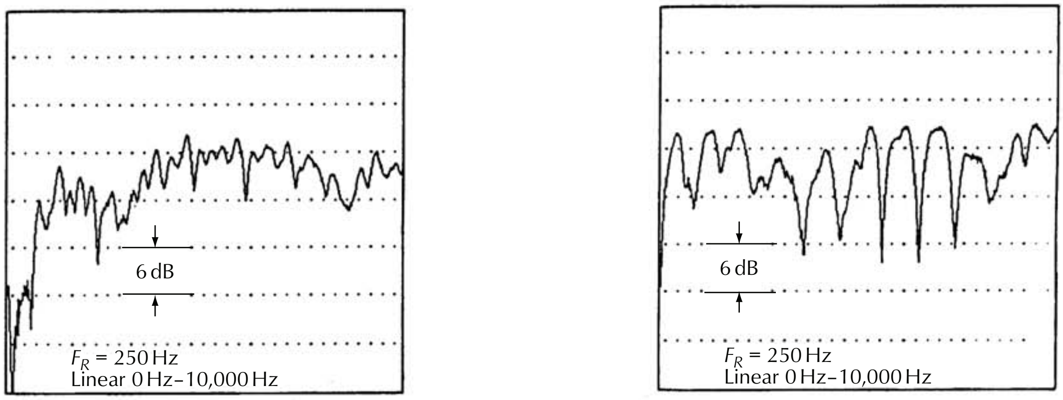
21.4.3 Polar Response

It was apparent that the missynchronization caused more than comb filtering of the signal on the measurement axis. We later duplicated the missynchronization by misaligning two loudspeakers by the same 4 in distance and measured the polar response. Fig. 21-10 shows the before and after synchronization ETCs. Fig. 21-11 shows the polar response before and after synchronization.

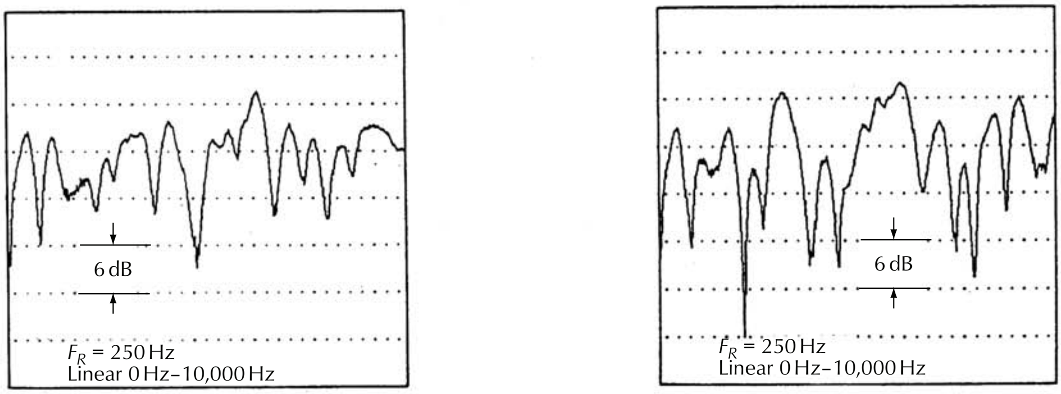
When a missynchronization as small as a few inches is present, it destroys a planned polar response. In this church, the polar response of the synchronized speakers encountered a usable *Ma* factor, which resulted in very little excitation of the reverberant sound field. The missynchronized array sent significant energy to the side walls and to the floor and ceiling, generating an audibly higher reverberant level. It also supplied energy via the ceiling to the rear wall where it was focused and returned to the measurement position. The change in energy ratio, $L_D - L_R$, was from −5.76 dB for the unsynchronized and +0.75 dB for the synchronized system.



A. Loudspeaker array configuration. B. Measurement of a single horn #7.



C. Measurement of high frequency horn #5. D. Measurement of high frequency horns #3 and #5.



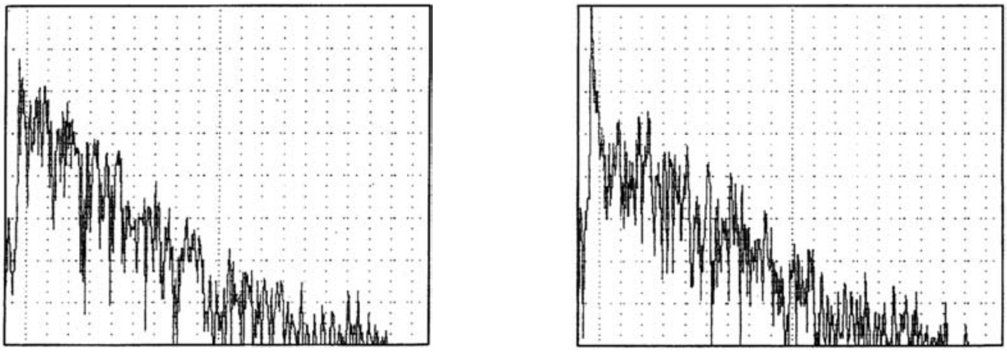
E. Measurement of high frequency horns #3, #5, and #7. F. Measurement of high frequency horns #3 through #8.

Figure 21-7. Effect of combining unaligned loudspeakers in an array.

21.5 Conclusion

A good, basic understanding of these simple but important relationships is vital to the creative design of sound systems today. Signal delay relationships complete the chain of necessary criteria for the proper placement of a loudspeaker.

1. Location allows an AL_{CONS} of 15% or less at D_2. Location allows $PAG = NAG$.
2. Location ensures that no signal delay shall exceed 40 ms (approximately 45 ft) at either the listener's location or the performer's location without sufficient remedial measures (earphones for the announcer at a basketball arena, monitor loudspeakers, or similar facilities).

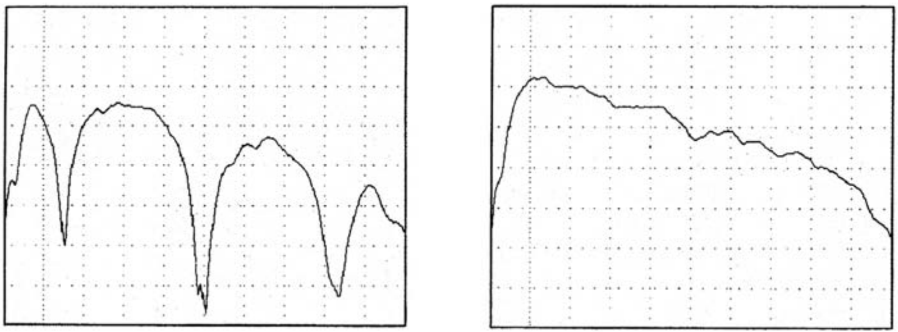

Vertical: 6 dB/division
Horizontal: 0 - 1,971,547 ms
No delay

A. Near-throw and far-throw horns out of alignment 3 inches (mouths of the horns were aligned).

Vertical: 6 dB/division
Horizontal: 0 - 1,971,547 ms
Both horns: Near -throw delayed 300 ms

B. Both horns synchronized with near-throw horn.

Figure 21-8. ETCs of a church sound system.

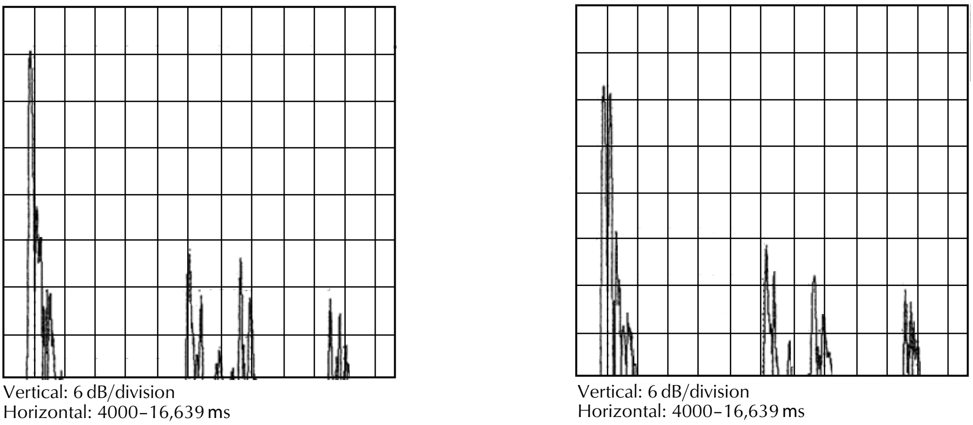

Vertical: 6 dB/division
Horizontal: 50.33 - 10,001.20 Hz
Resolution: 5.3674E+01 Hz
Both horns: No delay

A. Unaligned horns.

Vertical: 6 dB/division
Horizontal: 50.33 - 10,001.20 Hz
Resolution: 5.3674E+02 Hz
Both horns: Near -throw delayed 300 ms

B. Frequency response of synchronized horns.

Figure 21-9. Frequency response of unaligned and synchronized horns.

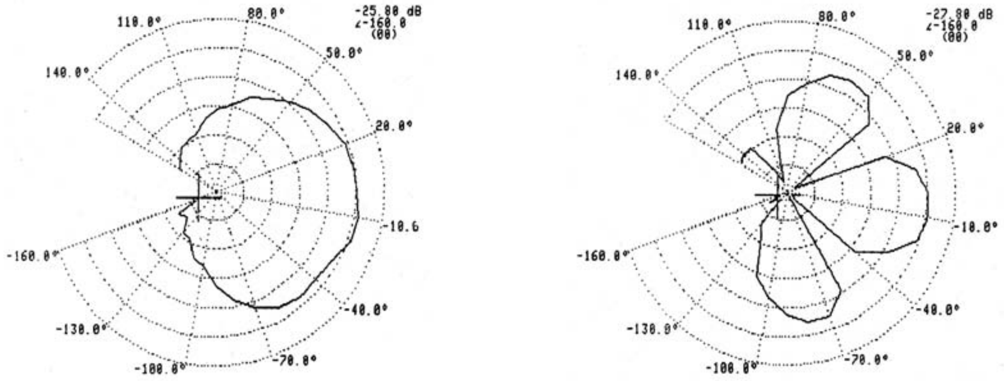

Vertical: 6 dB/division
Horizontal: 4000–16,639 ms

A. In alignment.

Vertical: 6 dB/division
Horizontal: 4000–16,639 ms

B. Out of alignment—note 6 dB drop in level.

Figure 21-10. Stacked monitors.



A. In alignment. B. Out of alignment.

Figure 21-11. Two monitors in and out of alignment.

The wise use of digital delay devices in audio systems required the development of low noise, reliable devices, precision measurement analyzers, and last, but certainly not least, a cadre of trained personnel capable of applying the tools and techniques developed over the years. Modern day tools, techniques, and engineers handle these challenges successfully.

Bibliography

H. Haas. “The Influence of a Single Echo on the Audibility of Speech,” *J. Audio Eng. Soc.*, Vol. 20 (Mar. 1972), pp. 145-159.

J. Henry. “On Acoustics Applied to Public Buildings.” Paper presented at the American Association for the Advancement of Science, August 1856; later published in the Smithsonian Institution publication.

J. K. Hilliard. “Notes on How Phase and Delay Distortions Affect the Quality of Speech, Music and Sound Effects,” *IEEE Trans. on Audio* (Mar.-Apr. 1964).

Chapter 22

Signal Processing

by Eugene Patronis, Jr.

22.1 Spectra . . . 477
  Fourier Trigonometric Series . . . 477
  Fourier Exponential Series . . . 481
  Fourier Integral . . . 483
  General Properties of Fourier Transforms . . . 484
  Unit Impulse . . . 487
  Spectrum of a Periodic Impulse Sampled Signal . . . 489
  Sampling Theorem . . . 493
  Aliasing . . . 494
  Realistic Sampling . . . 495
22.2 Analog to Digital Conversion . . . 497
22.3 System Theory . . . 502
  Pole and Zero Analysis . . . 508
  Further Considerations . . . 512
  Equalization . . . 514
  Equalization—Global or Local . . . 517
22.4 Digital Systems and the Z Transform . . . 520
  Recursive or IIR Filters . . . 524
  Linear Phase Filters . . . 527
22.5 Dynamics Processing . . . 528
Bibliography . . . 531

One could reasonably argue that anything interposed between the original source of sound and the listener constitutes signal processing. Many recording and sound reinforcement engineers, for example, select microphones based upon how the microphone alters the quality of the sound in the final product. Some microphones provide a boost in the 2 kHz to 4 kHz region and are said to add "presence." Cardioid microphones when worked at a close distance exhibit a "proximity effect" that amounts to a bass boost that can add body to a weak or thin voice.

For the present purpose, however, signal processing will be confined to certain properties of the signal chain that exist between the microphone or microphones and the loudspeaker or loudspeakers. Much of this processing is linear and time invariant in that it does not depend on the signal amplitude or the time of occurrence of the signal in question. In some instances the processing is non-linear such as noise gates, downward expanders, and compressors as well as limiters.

Amplifiers also exist in this signal chain. Wide bandwidth linear amplifiers, other than offering voltage and or power amplification, are essentially benign and as such are not considered as signal processors per se.

Linear signal processing will be considered to be any and all filtering that modifies a signal's amplitude and phase as a function of frequency. System equalization falls in this category. Linear signal processing will also include all pass filters that modify only phase as a function of frequency while leaving the amplitude untouched. All pass filters also include signal delay units. Signal delay units, while delaying the appearance of a signal as viewed on the time axis, accomplish this by providing a phase lag that is proportional to frequency while leaving the amplitude untouched.

Signal processing functions are often distributed among several devices in the overall signal chain. Mixers, for example, often contain several boost/cut bandpass or shelving filters associated with each input channel. Additionally, loudspeaker management systems generally offer loudspeaker crossover networks as well as signal delay for individual loudspeaker elements.

Originally, signal processing was accomplished solely through the employment of analog circuitry. Currently, both analog as well as digital circuitry are employed with more and more digital circuitry being introduced almost on a daily basis. In this regard, analog to digital and digital to analog converters are necessary adjuncts to the dedicated digital signal processors. These devices will be referred to as ADCs, DACs, and DSPs, respectively. The ADC employs a process known as signal sampling while the DAC employs a process termed signal reconstruction. The DSP can be programmed to perform digital filtering, signal delay, equalization, and other signal processing functions. In order to understand the operations involved in signal processing it is necessary to know how signals are described in the frequency domain. Furthermore, the interactions between the various components in a signal processing chain are best understood in terms of subject matter known as system theory. These two areas will be reviewed briefly before examining the details of signal processing itself.

22.1 Spectra

Jean Baptiste Joseph Fourier (1768-1830) was a French mathematician and theoretical physicist. While working on a problem dealing with heat conduction in solids, Fourier made a mathematical discovery which has had significance far beyond the bounds of his original problem. Fourier's original discovery has led to the development of the modern mathematical tools that allow a frequency domain description of events which occur in the time domain.

22.1.1 Fourier Trigonometric Series

In Fourier's original problem, a source of heat was placed at some point within a thin disc and the objective was to describe the temperature as a function of position on the circular periphery of the disc. The geometry of the disc suggested polar coordinates as the coordinates of choice for describing position both in the interior of the disc as well as on the periphery. The describing function for the temperature distribution on the periphery of the disc for this choice of coordinates is periodic in the polar angle with a period of 2π. This is true because starting with any initial angle θ, an increase of angle by 2π brings one back to the same point having, of course, the same temperature. Fourier found that upon letting y equal the temperature at a given point on the periphery that

$$\begin{aligned} y &= f(\theta) \\ &= \left(a_0 + \sum_{n=1}^{\infty} a_n \cos(n\theta) + b_n \sin(n\theta) \right) \end{aligned} \tag{22-1}$$

In the above expression, n takes on the value of each positive integer. The coefficient a_0 is the average value of y over the interval 0 to 2π.

$$a_0 = \langle y \rangle = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)\,d\theta \tag{22-2}$$

The coefficients in the infinite series sum representing *y* are given by

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(\theta)\cos(n\theta)\,d\theta \tag{22-3}$$

and

$$b_n = \frac{1}{\pi}\int_0^{2\pi} f(\theta)\sin(n\theta)\,d\theta \tag{22-4}$$

The expression for $f(\theta)$ is now called the Fourier trigonometric series. Pursuant to Fourier's work, it was soon discovered that the Fourier trigonometric series were members of a much larger set of mathematical functions, called orthogonal functions, having similar properties. The orthogonal properties of the members of the Fourier series are summarized in the following statements of fact where *m* and *n* are integers. When *m* and *n* are any integers,

$$\int_0^{2\pi} \cos(m\theta)\,d\theta = 0 \tag{22-5}$$

$$\int_0^{2\pi} \sin(m\theta)\,d\theta = 0 \tag{22-6}$$

$$\int_0^{2\pi} \sin(m\theta)\cos(n\theta)\,d\theta = 0 \tag{22-7}$$

In the event that *m* and *n* are different integers,

$$\int_0^{2\pi} \cos(m\theta)\cos(n\theta)\,d\theta = 0 \tag{22-8}$$

$$\int_0^{2\pi} \sin(m\theta)\sin(n\theta)\,d\theta = 0 \tag{22-9}$$

Finally, for the case of *m* and *n* being equal integers,

$$\int_0^{2\pi} \cos^2(m\theta)\,d\theta = \pi \tag{22-10}$$

$$\int_0^{2\pi} \sin^2(m\theta)\,d\theta = \pi \tag{22-11}$$

Now that the essential tools are in place, the foregoing can be made more meaningful by performing a Fourier analysis on some common periodic waveforms which are often encountered in audio and electronics. A useful waveform often employed in amplifier testing and evaluation is the square wave. The square wave to be considered here has an amplitude of 1 V and a period of *T*. This square wave is depicted in Fig. 22-1.

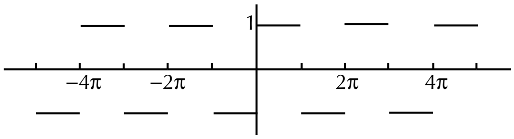


Figure 22-1. A square wave of unit amplitude.

The depiction in Fig. 22-1 employs the independent variable θ along the horizontal axis. The independent variable can easily be converted to time by recognizing that

$$\theta = 2\pi\frac{t}{T} = 2\pi f_0 t = \omega_0 t \tag{22-12}$$

The objective at this point is to determine the terms in the Fourier trigonometric series which when added together will produce the shape depicted in Fig. 22-1. Initially, two observations should be made about the waveform in the figure. The first observation is that the average value of the waveform over any period or any integral number of periods is zero. Thus any collection of terms employed to describe the waveform must also have an average value of zero. The second observation is that

$$f(\theta) = -f(-\theta) \tag{22-13}$$

In words, this says that the algebraic sign of the *y* axis value or ordinate of the waveform reverses as one goes from a positive point on the *x* axis or abscissa of the waveform to the corresponding negative value of the abscissa. Mathematically, such a function is called an odd function and can only be represented by a collection of odd functions. An even function would be one for which

$$f(\theta) = f(-\theta) \tag{22-14}$$

These observations save a lot of unnecessary labor because the terms in the Fourier series are either sines or cosines. The sine is an odd function while the cosine is an even function. At this point, then, only the sine terms in the Fourier series need be considered for this particular waveform.

The example square wave has a rather simple mathematical description, namely

$$f(\theta) = +1 \text{ when } 0 < \theta < \pi \tag{22-15}$$

$$f(\theta) = -1 \text{ when } \pi < \theta < 2\pi \tag{22-16}$$

At this juncture it is possible to set up a general expression for the coefficients of the sine terms in the Fourier series describing the example square wave.

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_0^{\pi}(+1)\sin(n\theta)d\theta + \frac{1}{\pi}\int_{\pi}^{2\pi}(-1)\sin(n\theta)d\theta \\ &= \frac{1}{n\pi}[-\cos(n\pi) + 1 + \cos(2n\pi) - \cos(n\pi)] \\ &= \frac{1}{n\pi}[1 - 2\cos(n\pi) + \cos(2n\pi)] \end{aligned} \tag{22-17}$$

One is now positioned to write the Fourier series describing the given square wave, i.e.; specific numbers can be assigned to the coefficients. As mentioned earlier, $a_n = 0$ for all values of n. $b_0 = 0$ because the integrands in the general expression vanish as $sin(0)$ is zero. For even values of n other than zero

$$\begin{aligned} b_n &= \frac{1}{n\pi}[1 - 2 + 1] \\ &= 0 \end{aligned} \tag{22-18}$$

where,
n is an even integer.

For odd values of n

$$\begin{aligned} b_n &= \frac{1}{n\pi}[1 + 2 + 1] \\ &= \frac{4}{n\pi} \end{aligned} \tag{22-19}$$

where,
n is an odd integer.

Therefore the unit amplitude square wave under consideration is represented by the infinite series

$$\begin{aligned} f(\theta) = \frac{4}{\pi}\Big[&\sin(\theta) + \frac{1}{3}\sin(3\theta) + \frac{1}{5}\sin(5\theta) \\ &+ \frac{1}{7}\sin(7\theta) + \ldots\Big] \end{aligned} \tag{22-20}$$

Recalling that $\theta = \omega_0 t = 2\pi f_0 t$, this result can also be expressed as

$$\begin{aligned} f(t) = \frac{4}{\pi}\Big[&\sin(2\pi f_0 t) + \frac{1}{3}\sin(2\pi 3 f_0 t) \\ &+ \frac{1}{5}\sin(2\pi 5 f_0 t) + \frac{1}{7}\sin(2\pi 7 f_0 t) + \ldots\Big] \end{aligned} \tag{22-21}$$

The spectrum of the square wave, then, consists of a fundamental frequency equal to the reciprocal of the period of the square wave along with diminishing amplitude odd harmonics of the fundamental frequency. Furthermore, as the phase angles of the fundamental as well as the harmonics are zero at $t = 0$, all of the frequency components are in phase.

It is instructive to examine how the Fourier description of the square wave is affected by the choice of origin of coordinates. Fig. 22-2 displays the unit amplitude square wave where the origin of coordinates has been shifted to the right by $\pi/2$ radians. This new square wave is now an even function.

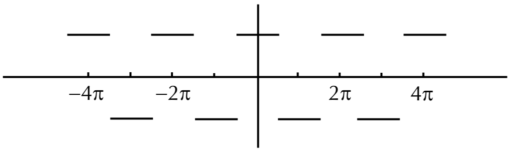


Figure 22-2. Unit amplitude square wave as an even function.

Mathematically, this shift amounts to a change of independent variable such that $2\pi f_0 t$ is replaced by $2\pi f_0 t' + \pi/2$. Upon making this substitution, $f(t)$ becomes $f(t')$ with

$$\begin{aligned} f(t') = \frac{4}{\pi}\Big[&\sin\left(2\pi f_0 t' + \frac{\pi}{2}\right) + \frac{1}{3}\sin\left(2\pi 3 f_0 t' + \frac{3\pi}{2}\right) \\ &+ \frac{1}{5}\sin\left(2\pi 5 f_0 t' + \frac{5\pi}{2}\right) + \ldots\Big] \end{aligned} \tag{22-22}$$

This can be greatly simplified by making use of a common trigonometric identity involving the sum of two angles.

$$\sin(A+B) = \sin(A)\cos(B) + \cos(A)\sin(B) \tag{22-23}$$

This identity leads to the statements

$$\sin\left(2\pi f_0 t' + \frac{\pi}{2}\right) = +\cos(2\pi f_0 t') \tag{22-24}$$

$$\sin\left(2\pi 3 f_0 t' + \frac{3\pi}{2}\right) = -\cos(2\pi 3 f_0 t') \tag{22-25}$$

$$\sin\left(2\pi 5 f_0 t' + \frac{5\pi}{2}\right) = +\cos(2\pi 5 f_0 t') \tag{22-26}$$

Therefore the expression for this time shifted square wave that is now an even function is

$$f(t') = \frac{4}{\pi}\left[\cos(2\pi f_0 t') - \frac{1}{3}\cos(2\pi 3 f_0 t') + \frac{1}{5}\cos(2\pi 5 f_0 t') - \frac{1}{7}\cos(2\pi 7 f_0 t') + \ldots\right] \tag{22-27}$$

It is important to observe that the time shift alters only the phases of the frequency components. The frequency components present and their amplitudes are unchanged.

There are some limitations on the types of mathematical functions which can be represented by the Fourier series. For example, the function to be represented in a given interval must be continuous and finite or, if discontinuous, must have a finite number of finite discontinuities. The mathematical square wave used as the first example has a discontinuous jump between +1 and −1 at $\theta = \pi$ and another discontinuous jump between −1 and +1 at $\theta = 2\pi$. At such points, the sum of the infinite series converges to the mid point of the jump. In this instance, the mid point is zero. Any physically generated square wave would make this transition in a finite though short interval and thus would not display discontinuous behavior. The Fourier series can represent such a physical square wave exactly at all points in the interval $0 \le \theta \le 2\pi$.

Before leaving the square wave expressed as an even function, one final change is of interest. In this change, a constant of 1 is added to the wave producing the waveform depicted in Fig. 22-3.

The square wave depicted in Fig. 22-3 has an average value of one. This follows from the fact that it was constructed by adding a constant value of one to a function which originally had an average value of zero. The formal calculation proceeds through the following steps.

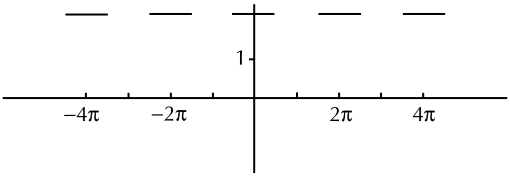


Figure 22-3. Square wave whose average value is one.

$$\begin{aligned} a_0 &= \frac{1}{2\pi}\int_0^{2\pi} f(\theta)d\theta \\ &= \frac{1}{2\pi}\left[\int_0^{\frac{\pi}{2}} 2d\theta + \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} (0\cdot d\theta) + \int_{\frac{3\pi}{2}}^{2\pi} 2d\theta\right] \\ &= \frac{1}{2\pi}[\pi + 0 + \pi] = 1 \end{aligned} \tag{22-28}$$

Upon equating θ with $2\pi f_0 t$, the Fourier series describing the square wave of Fig. 22-3 becomes

$$f(t) = 1 + \frac{4}{\pi}\left[\cos(2\pi f_0 t) - \frac{1}{3}\cos(2\pi 3 f_0 t) + \frac{1}{5}\cos(2\pi 5 f_0 t) - \frac{1}{7}\cos(2\pi 7 f_0 t) + \ldots\right] \tag{22-29}$$

An analysis of a half wave rectified cosine waveform of unit amplitude leads to the result

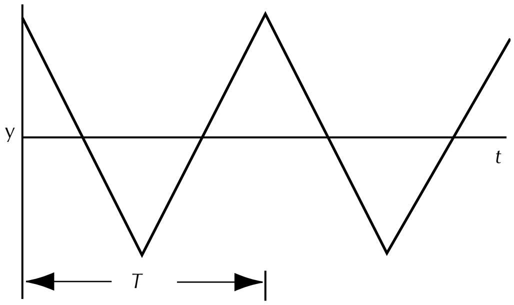
$$y(t) = \frac{1}{\pi}\left(1 + \frac{\pi}{2}\cos(2\pi f_0 t) + \frac{2}{3}\cos(2\pi 2 f_0 t) - \frac{2}{15}\cos(2\pi 4 f_0 t) + \frac{2}{35}\cos(2\pi 6 f_0 t) - \ldots\right) \tag{22-30}$$

If, instead, one has a full wave rectified cosine wave of unit amplitude, the Fourier analysis yields

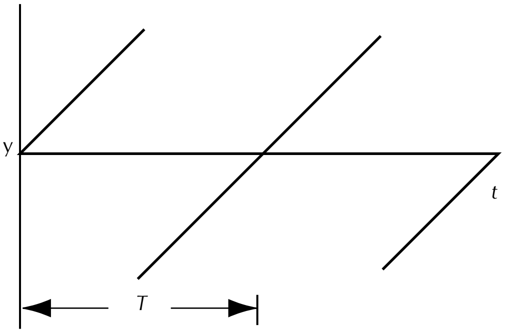
$$y(t) = \frac{2}{\pi}\left(1 + \frac{2}{3}\cos(2\pi 2 f_0 t) - \frac{2}{15}\cos(2\pi 4 f_0 t) + \frac{2}{35}\cos(2\pi 6 f_0 t) - \ldots\right) \tag{22-31}$$

It should be noted that the average value of the full wave signal is twice that of the half wave as one might guess. Additionally, the lowest frequency component in the full wave signal is the second harmonic rather than the fundamental frequency. Thus the full wave signal is more readily filtered than is the half wave.

Two other often-encountered waveforms are the triangle and the sawtooth. A plot of each of these waveforms is displayed in Fig. 22-4.



Triangle



Sawtooth

Figure 22-4. Triangle and sawtooth waveforms.

For the triangle waveform of unit amplitude

$$y(t) = \frac{8}{\pi^2}\left(\cos(2\pi f_0 t) + \frac{1}{9}\cos(2\pi 3 f_0 t) + \frac{1}{25}\cos(2\pi 5 f_0 t) + \ldots\right) \tag{22-32}$$

Whereas for the sawtooth of unit amplitude

$$f(t) = \frac{2}{\pi}\left[\sin(2\pi f_0 t) - \frac{1}{2}\sin(2\pi 2 f_0 t) + \frac{1}{3}\sin(2\pi 3 f_0 t) - \frac{1}{4}\sin(2\pi 4 f_0 t) + \ldots\right] \tag{22-33}$$

22.1.2 Fourier Exponential Series

Upon identifying $\omega_0 = 2\pi(1/T) = 2\pi f_0$, the general Fourier trigonometric series can be written as

$$f(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n\omega_0 t) + b_n \sin(n\omega_0 t) \tag{22-34}$$

If one writes Euler's theorem in the form $e^{jn\omega_0 t} = \cos(n\omega_0 t) + j\sin(n\omega_0 t)$ and solves individually for the sine and cosine terms, one obtains

$$\cos(n\omega_0 t) = \frac{1}{2}(e^{jn\omega_0 t} + e^{-jn\omega_0 t}) \tag{22-35}$$

$$\sin(n\omega_0 t) = \frac{-j}{2}(e^{jn\omega_0 t} - e^{-jn\omega_0 t}) \tag{22-36}$$

When these are substituted into the general expression for the trigonometric series, the result can be written as

$$f(t) = \sum_{n=-\infty}^{n=+\infty} c_n e^{jn\omega_0 t} \tag{22-37}$$

This compact expression is called the Fourier exponential series. The coefficients, cn, are related to the former coefficients in the following manner.

$$c_0 = a_0 \tag{22-38}$$

$$c_1 = \frac{1}{2}(a_1 - jb_1) \tag{22-39}$$

$$c_{-1} = \frac{1}{2}(a_1 + jb_1) \tag{22-40}$$

...

It should be noted that the coefficients for the exponential series, with the exception of c_0, are complex and that they occur in complex conjugate pairs such that

$$c_{-n} = c_n^* \tag{22-41}$$

The coefficients for the exponential series can be calculated directly by employing the integral expression

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-jn\omega_0 t}\, d(\omega_0 t) \tag{22-42}$$

The Fourier exponential series is introduced here for two reasons. Firstly, the general expression for the exponential series is more compact than that for the trigonometric series and, secondly, the exponential series for periodic waveforms provides a stepping stone to the Fourier integral transform. The Fourier integral transform can be employed to calculate the spectra of more general time dependent functions independent of whether the time function is periodic or not.

The Fourier exponential series will now be employed to determine the spectrum associated with a recurrent train of pulses. This periodic waveform is depicted in Fig. 22-5.

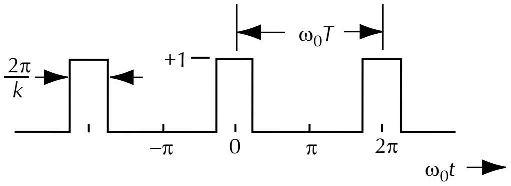


Figure 22-5. Recurrent pulse train.

The recurrent pulse train of Fig. 22-5 has an amplitude of 1, a fundamental frequency or pulse repetition frequency f_0 corresponding to the period T, and the ratio of the pulse repetition period to the duration of each pulse is denoted by a constant termed k. The figure was constructed assuming that k was equal to 4. The constant, k, however, can take on any value greater than one. If k is allowed to take on larger and larger values, the width of an individual pulse becomes narrower and narrower and the interval between pulses becomes relatively larger and larger.

The analysis of this recurrent pulse train consists of calculating the coefficients of the Fourier exponential series. In performing this analysis, the 2π interval over which the integrations are performed may be taken anywhere on the θ axis. For ease of calculation, this interval is chosen as $-\pi$ to $+\pi$. The function, however, is zero except between $-\pi/k$ and $+\pi/k$ in which range it has the value of one. Therefore when it is remembered that $\omega_0 t = \theta$, then

$$c_n = \frac{1}{2\pi}\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} e^{-jn\theta}\,d\theta \qquad (22\text{-}43)$$

If $n = 0$,

$$c_0 = \frac{1}{2\pi}\left(\frac{\pi}{k}+\frac{\pi}{k}\right) = \frac{1}{k} \qquad (22\text{-}44)$$

If $n \neq 0$,

$$\begin{aligned} c_n &= \frac{1}{-jn2\pi}\left(e^{-jn\frac{\pi}{k}} - e^{jn\frac{\pi}{k}}\right) \\ &= \frac{1}{n\pi}\times\frac{e^{jn\frac{\pi}{k}} - e^{-jn\frac{\pi}{k}}}{2j} \\ &= \frac{1}{k}\times\frac{\sin\left(n\frac{\pi}{k}\right)}{n\frac{\pi}{k}} \end{aligned} \qquad (22\text{-}45)$$

The coefficients, c_n, represent the amplitude and phase of the constituent frequency components which contribute to the recurrent pulse train. The recurrent pulse train is recovered or synthesized through the addition of these frequency components employing the respective amplitudes and phases determined in the analysis. Therefore,

$$f(t) = \sum_{n=-\infty}^{n=\infty} \frac{1}{k}\times\frac{\sin\left(n\frac{\pi}{k}\right)}{n\frac{\pi}{k}}e^{jn\omega_0 t} \qquad (22\text{-}46)$$

The fact that n takes on both positive and negative values suggests the mathematical necessity for considering both positive and negative frequencies, i.e., $\pm n\omega_0$. Fig. 22-6 is a plot of c_n versus $n\omega_0$ in the instance where k has the value two. This assignment forces the recurrent pulse train into being a square wave with an average value of one-half. Such a plot is termed a Fourier line spectrum as the frequencies are discrete.

The line spectrum of the square wave recurrent pulse displayed in Fig. 22-6 exhibits the average value or dc term at zero frequency, the fundamental components at $\pm\omega_0$, and the characteristic odd harmonics with alternating signs and diminishing amplitudes according to the harmonic order. Nothing is displayed beyond the fifth harmonic because of diminishing scale even though all odd harmonics are present in the actual spectrum.

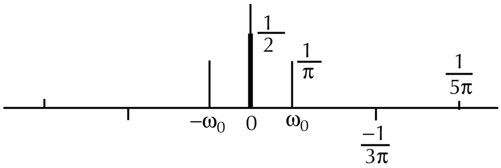


Figure 22-6. Fourier line spectrum of recurrent pulse as a square wave.

22.1.3 Fourier Integral

Most of the signals of interest in acoustics and electronics are dynamic signals which do not repeat themselves over and over. In other words, such signals are not periodic yet it is still necessary, in fact essential, to know the spectral content of such signals. One way to approach the analysis of signals that do not repeat themselves anywhere on the time axis is to consider that the period of such signals is infinite. To that end, the expressions for the analysis and synthesis employing the Fourier exponential series are recast in a form where time is taken as the independent variable and the period appears explicitly. These new expressions appear as Eq. 22-47 and Eq. 22-48.

$$c_n = \frac{1}{T}\int_{\frac{-T}{2}}^{\frac{T}{2}} f(t)e^{-j\omega_n t}dt \qquad (22\text{-}47)$$

$$f(t) = \sum_{n=-\infty}^{n=\infty} c_n e^{j\omega_n t} \qquad (22\text{-}48)$$

where,

$$\omega_n = n\omega_0, \qquad (22\text{-}49)$$

$$T = \frac{2\pi}{\omega_0}. \qquad (22\text{-}50)$$

At first glance, we might conclude that if T is allowed to approach infinity in Eq.22-47, that c_n approaches zero and all is lost. A way out is provided if we also observe in Eq. 22-50 that as T approaches infinity, ω_0 must be shrinking and approaching zero. In the integral expression of Eq. 22-47, if we now substitute for T from Eq. 22-50 and then divide both sides of the expression by $\omega_0/2\pi$, we obtain

$$\frac{c_n}{\frac{\omega_0}{2\pi}} \equiv F_n = \int_{\frac{-T}{2}}^{\frac{T}{2}} f(t)e^{-j\omega_n t}dt \qquad (22\text{-}51)$$

As long as $f(t)$ is reasonably well behaved such that the integral does not diverge, the quantity F_n will contain the information which is sought. The corresponding synthesis expression can now be written as

$$f(t) = \frac{1}{2\pi}\sum_{n=-\infty}^{n=+\infty} F_n e^{j\omega_n t}\omega_0 \qquad (22\text{-}52)$$

At this juncture, it is possible to determine what happens when the period grows indefinitely large such that in the limit it becomes infinite. As T grows larger and larger, ω_0 grows smaller and smaller and it becomes an infinitesimal called $d\omega$. As ω_0 grows small, the harmonic frequencies designated by ω_n become so numerous and closely packed that they become a continuum represented by the continuous variable called ω. Additionally, instead of having a discrete set of values F_n for each harmonic, one now has a continuous function of frequency designated by $F(\omega)$. Finally, in the limit of infinite T, the infinite sum of discrete terms of Eq. 22-52 becomes a continuous sum or integral. Finally then, Eqs. 22-51 and 22-52 become

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt \qquad (22\text{-}53)$$

and

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{j\omega t}d\omega \qquad (22\text{-}54)$$

Eqs. 22-53 and 22-54 constitute the Fourier transform and inverse Fourier transform, respectively, applicable to time functions whether they are periodic or not.

Not every conceivable $f(t)$ has a Fourier transform, i.e., the integral indicated in Eq. 22-53 may not be calculable. Fortunately, for most of the signals encountered in acoustics and communications, this is not the case and $F(\omega)$ can be calculated. One of the many useful properties possessed by a Fourier transform is that of uniqueness. This means that if we can calculate the Fourier transform of a given $f(t)$ there is one and only one $F(\omega)$ associated with this $f(t)$. In fact the $F(\omega)$ so calculated embodies all of the information contained in $f(t)$ expressed in a different way. In the case of $f(t)$, the independent variable is time which is the variable closest to the nature of human experience. In the case of $F(\omega)$, the independent variable is angular frequency and the domain of description is called the frequency or spectral domain. Time and frequency constitute two different descriptors of the same physical phenomena. They are both useful in acquiring a deeper understanding of the nature of physical systems. In fact, the keys to the understanding of the nature of atomic structure were extracted almost solely by studying the spectra of the radiation emitted by excited atoms.

22.1.4 General Properties of Fourier Transforms

There are some properties common to all Fourier transforms, knowledge of which can save much labor in transform calculations. Physical signals are generally real functions of time but the Fourier transforms of such signals may be real or complex depending on the mathematical structure of the time signal. For example, if the time signal is real and an even mathematical function of t, the Fourier transform is also real and an even mathematical function of ω. If $f(t)$ is real and odd, then $F(\omega)$ is imaginary and odd. If $f(t)$ is real and neither even nor odd, then $F(\omega)$ is complex.

A second useful property is that of time shift. Suppose it is desired to shift the position of a given time function $f(t)$ on the time axis for a fixed amount of time t_0 such that $f(t)$ is replaced by $f(t-t_0)$ indicating the time function is shifted to the right by an amount of time equal to t_0 while the shape of the function remains unchanged. Then if $F(\omega)$ is the transform of $f(t)$, the transform of $f(t-t_0)$ is $e^{(-j\omega)t_0}F(\omega)$. A corollary to this property is that of frequency shift. Suppose that the transform of $f(t)$ is again $F(\omega)$. If one shifts the transform to the right along the frequency axis by a fixed amount ω_0 then $F(\omega)$ becomes $F(\omega-\omega_0)$ and the time function of which this is the transform becomes $e^{j\omega_0 t}f(t)$. A shift to the left would be accomplished by a change in the algebraic sign of the fixed quantity in both instances.

Additionally, the process of taking the Fourier transform is a linear operation which allows one to apply the principle of superposition. Suppose there are two different time functions denoted as $f(t)$ and $g(t)$ which have transforms denoted as $F(\omega)$ and $G(\omega)$ respectively. From $f(t)$ and $g(t)$ form a new time function given by $af(t)+bg(t)$ where a and b are constants. As a result of the linearity property, the transform of this new time function will be $aF(\omega)+bG(\omega)$.

Another useful property is associated with a mathematical process known as convolution. This process is a little trickier to describe as well as to actually perform. The convolution process can be performed in either the frequency or time domains. In order to gain an understanding of the nature of the process an example will be drawn from the field of spectrum measurement. Suppose it is desired to determine the shape of the spectrum produced by some stationary signal source. Determining the shape of the spectrum amounts to continuously identifying the values of the frequencies which are present as well as the strength of the signal at each value of the frequency. A stationary signal source would be one whose spectrum does not change with time and hence the measurement may be carried out in a leisurely fashion.

One way of performing the measurement would be to apply the signal to a tunable bandpass filter and to plot the root mean square value of the output of the filter as it is slowly tuned over the complete range of all frequencies of interest. The shape of the plotted results will depend not only on the actual shape of the spectrum to be measured but also on the shape of the tunable filter's frequency response. For example, at each frequency of tuning, the amount of signal passed by the filter obviously depends on whether the filter's response curve is narrow or broad, i.e., whether the filter Q is large or small. The shape of the measured curve will be that of the convolution of the actual spectrum with that of the filter's response function.

Convolution can be applied in either the time or frequency domain. A detailed calculation will now be presented in the time domain where time signals of simple though reasonable geometric shapes can be chosen so as to visualize easily what is being done at each step of the process. Initially a mathematical statement of the convolution process must be made. Suppose one has two time functions denoted as $f(t)$ and $g(t)$. The convolution of $f(t)$ with $g(t)$ is denoted and defined as

$$f(t) * g(t) \equiv \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau \tag{22-55}$$

In Eq. 22-55, τ is a dummy variable having the dimensions of time, which is integrated out producing a result that is a function of t only. The corresponding statement in the frequency domain might appear as

$$F(\omega) * G(\omega) \equiv \int_{-\infty}^{\infty} F(k)G(\omega-k)dk \tag{22-56}$$

In Eq. 22-56 k is the dummy variable that is integrated out in the convolution process.

Fig. 22-7 depicts the shape of the two time functions to be convolved in a geometrical illustration of the convolution process.

The first step in describing the convolution process geometrically is to flip the figure describing $g(t)$ about the vertical axis. This figure will now describe $g(t-\tau)$. It is the negative τ that dictates the

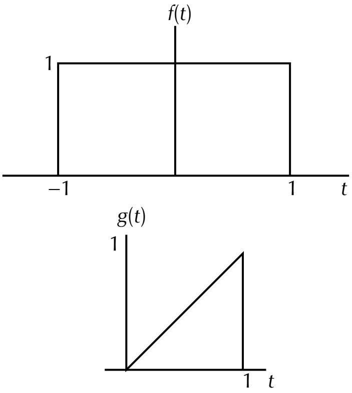


Figure 22-7. Functions to be convolved.

reversal. Next one positions the figure describing the *g* function sequentially for all values of *t*. At each value of *t*, the area where the two figures overlap is calculated. This area of overlap is the result of the integration over the dummy variable τ in the mathematical definition of convolution. Values of *t* that do not produce an overlap can be ignored. This is illustrated for several values of *t* in Fig. 22-8.

For *t* ranging between 0 and 1, the triangle will be completely within the rectangle and the overlap area will remain at a constant value of 0.5. The remaining steps of the process are illustrated in Fig. 22-9.

The result of the convolution calculation is a new function of time denoted by *h*(*t*). Mathematically,

$$h(t) = f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau \qquad (22\text{-}57)$$

For each value of *t*, the function *h*(*t*) has as its value, the value of the area of overlap of the two functions which are being convolved. The function *h*(*t*) for the present example is drawn in Fig. 22-10.

The property of the Fourier transform which makes the convolution of interest is that convolution in the time domain is equivalent to multiplication in the frequency domain. The meaning of this is as follows. Let $F(\omega)$ be the transform of $f(t)$ and $G(\omega)$ be the transform of $g(t)$. This being the case, then the Fourier transform of the convolution of $f(t)$ with $g(t)$ is simply $F(\omega)$ multiplied by $G(\omega)$. The converse is also true. If we were to convolve $F(\omega)$ with $G(\omega)$ in the frequency domain and then take the inverse Fourier transform of the result, the resulting calculation would be equal to the product of $f(t)$ with $g(t)$ in the time domain.

There is an important theorem associated with the Fourier transform which further illustrates the equality between the time domain and frequency domain descriptions. This is called Parseval's Theorem even though it was first expressed and used by Lord Rayleigh. This is called an energy theorem because it involves the integral with respect to time of the square of a signal property such as acoustic pressure. The acoustic power is proportional to the square of the acoustic pressure and the time integral of power leads to an expression of the signal energy. The formal statement of the theorem appears as Eq. 22-58.

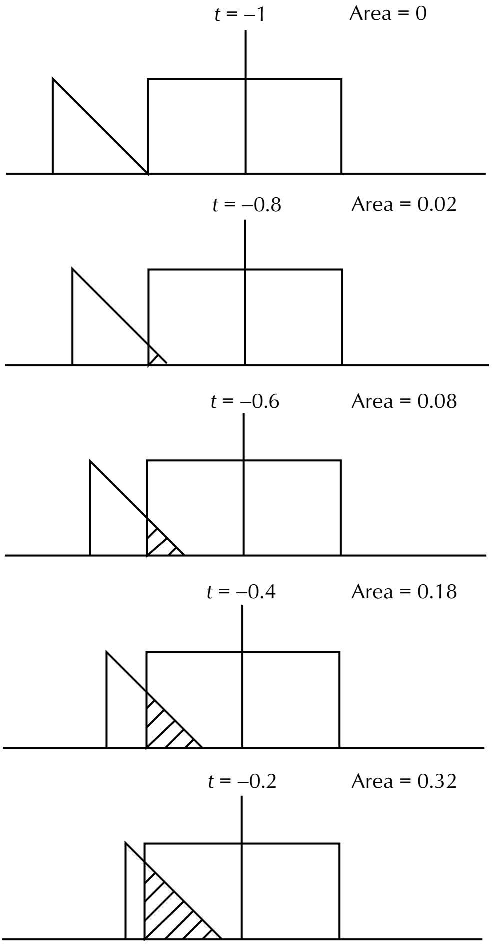


Figure 22-8. Early steps in the convolution process.

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2 d\omega \qquad (22\text{-}58)$$

In Eq. 22-58, the square of the absolute magnitude appears rather than simply just the square because even though $f(t)$ is real for physical signals and the square of the absolute magnitude and the simple square are equal for this function, $F(\omega)$ is often complex and its simple square would be complex whereas the square of its absolute

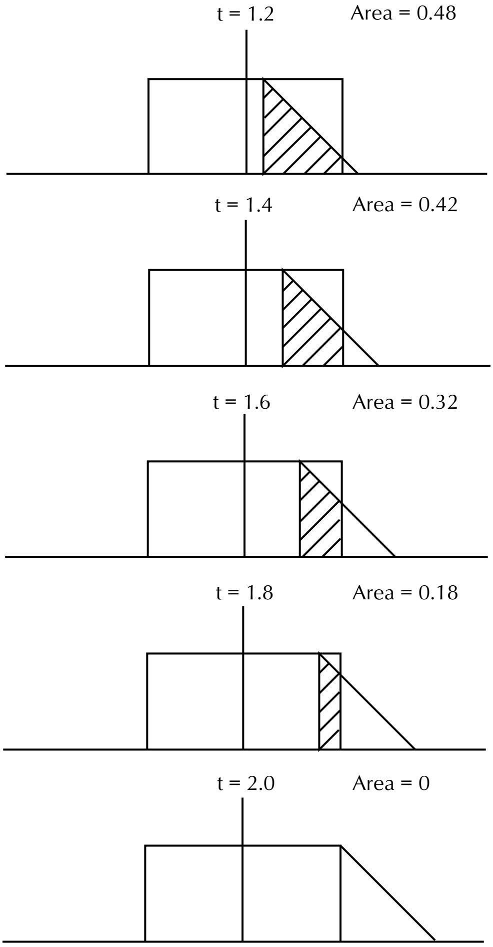


Figure 22-9. Final steps in the geometrical calculation of the convolution.

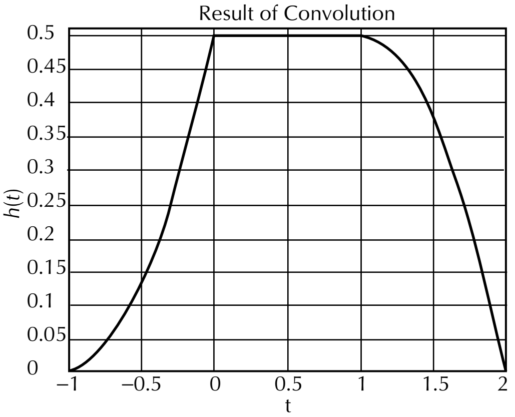


Figure 22-10. Convolution of $f(t)$ with $g(t)$.

magnitude is real and the result of this calculation must always be real. It is often much easier to perform the integral on the right rather than that on the left and hence the utility of the theorem for finite energy signals. The term $|F(\omega)|^2$ in Eq. 22-58 is called the spectral density function in that it represents the energy per unit angular frequency.

Another significant property of the time and frequency domain descriptions of a given signal becomes apparent when one compares the duration of a signal in the time domain with the frequency interval occupied by the same signal expressed in the frequency domain. In exploring this feature, an example will be drawn from the examination of a tone burst, which is a significant test signal employed in both acoustics and electroacoustics. A tone burst is formed by gating on a continuously operating sinusoidal oscillator at a fixed frequency for a time interval equal to an integral number of periods of the oscillator signal. Fig. 22-11 displays for an interval of one period the function $|f(t)|^2$ from Eq. 22-58 appropriate for a 1 kHz tone burst. Fig. 22-12 displays the corresponding frequency domain spectral density function for this tone burst.

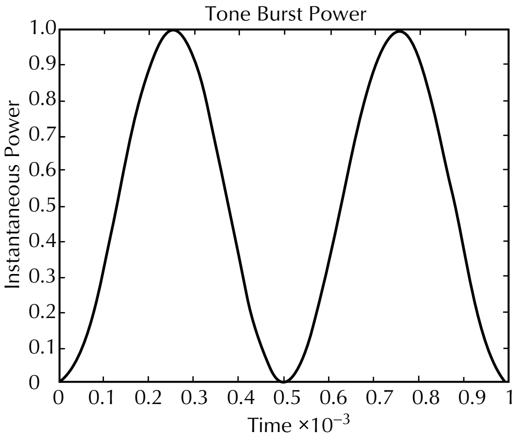


Figure 22-11. 1 kHz tone burst for one period.

It should be apparent from the figures that even though the burst's duration in the time domain is well defined and limited to an interval of 1 ms, the same cannot be said of the spectral density function. The spectral density function is broadly spread out over a range of frequencies. In fact, the positive and negative frequency peaks of the spectral density do not even occur at 1 kHz but rather at lower values. Note the positions of the markers at ±1 kHz. Contrast Fig. 22-12 with Fig. 22-13 which corresponds to a tone burst of not one period but rather of ten periods, that is, a duration in time larger by a factor of ten.

Now the spectral density function is much better defined and the peaks occur almost precisely at ±1 kHz as evidenced by the markers. In addition to being much narrower, the peaks are now much taller

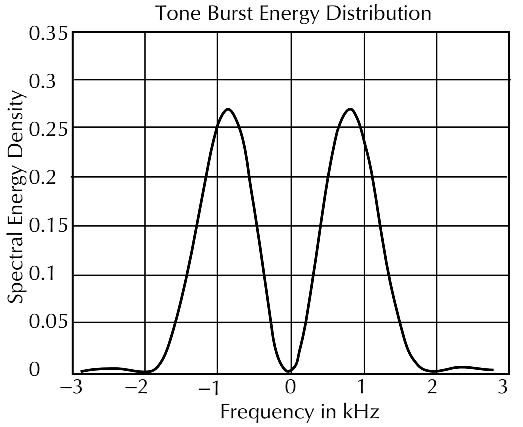


Figure 22-12. Spectral density function for a one period burst of 1 kHz.

as now the burst has delivered ten times the energy of the former case and the area under the total curve must also be ten times larger. This behavior results from another general property of the time and frequency domain descriptions of signals. This property is sometimes referred to as the classical uncertainty principle. Mathematically this can be expressed as

$$\Delta t \Delta f \geq 1 \tag{22-59}$$

The conclusion to be reached which is expressed in this relationship is that if the frequency content of a signal is to appear with precision, then the existence of the signal and our observation of it must endure for a long period of time. The converse is also true in that if a signal exists for only a brief time, a precise statement or measurement cannot be made about its frequency content.

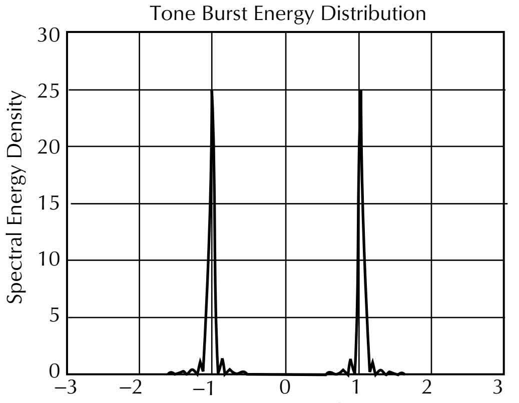


Figure 22-13. Spectral density function for a ten period burst at 1 kHz.

22.1.5 Unit Impulse

Ironically, oftentimes in physics and mathematics progress in thinking about things that actually exist in nature can be facilitated by thinking about things that do not. A case in point is that of the unit impulse. The physical thinking that led to the concept of the unit impulse will be discussed in the section on system theory. For the present purposes, we will be satisfied initially in describing the unit impulse by listing its mathematical properties. Firstly, the unit impulse is denoted as $\delta(t)$ where

$$\begin{aligned} \delta(t) &= 0, \text{ when } (t \neq 0) \\ \delta(t) &= \infty, \text{ when } (t = 0) \end{aligned} \tag{22-60}$$

Additionally, $\delta(t)$ is such that

$$\int_{-\infty}^{+\infty} \delta(t) dt = 1 \tag{22-61}$$

A corollary to Eq. 22-61 issues from the fact that $\delta(t)$ vanishes everywhere except at the origin therefore

$$\int_{a<0}^{b>0} \delta(t) dt = 1 \tag{22-62}$$

Eq. 22-62 is true if and only if a is taken to be any real number less than zero and b is taken to be any real number greater than zero. One becomes more comfortable with the above statements when one considers the following. Imagine a rectangle centered on the origin. This rectangle has a duration along the time axis of W and a height along the vertical or ordinate axis of H. Now, further imagine that H always equals to $1/W$. Clearly, the area under the rectangle being the product of H with W is unity. Note also that this area is dimensionless. Now, consider that the duration along the time axis, W, becomes progressively smaller. When this occurs, the height, H, becomes progressively larger in such a way that the product of H with W remains constantly at unity. In the limit as W goes to zero, H goes to infinity with the product of H with W remaining at unity. This limiting situation describes the unit impulse $\delta(t)$. A quantity that is zero everywhere except when the number in the parentheses, in this instance t, is zero. At the point where the number in the parentheses is zero, the impulse has an infinite value while retaining a total area of one. The area, of course, being given by the integral expression of Eq. 22-61 or Eq. 22-62.

Unit impulses may occur at points other than at the origin. For example, suppose one needs to describe a unit impulse that occurs at $t = t_0$. This is stated as $\delta(t - t_0)$. The impulse is located at the position where the number in the parentheses becomes zero. In this instance, this occurs where $t = t_0$. Also, for this impulse location, Eq.22-62 becomes

$$\int_{t<t_0}^{t>t_0} \delta(t - t_0)dt = 1 \qquad (22\text{-}63)$$

One could in fact conceive of a function of time denoted as $s(t)$ that represents an infinite repetitive pulse train of unit impulses. Let the repetition period of this train be denoted as T_s. The expression for $s(t)$ is actually fairly simple.

$$s(t) = \sum_{n=-\infty}^{n=+\infty} \delta(t - nT_s) \qquad (22\text{-}64)$$

where,

n is any integer.

For obvious reasons, unit impulses cannot be represented on drawings made to scale. For graphical purposes, bold arrows positioned appropriately on the abscissa axis represent unit impulses. The three cases just discussed might appear then as displayed in Fig. 22-14.

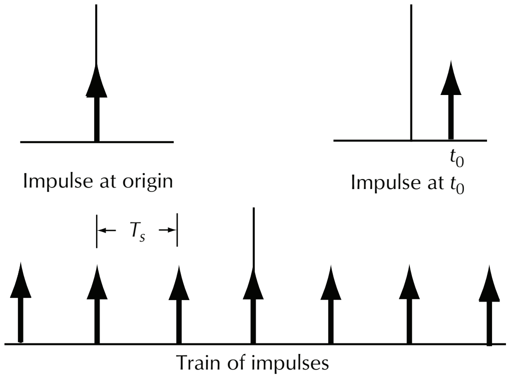


Figure 22-14. Various unit impulse depictions.

Now that the mathematical formalism is in place, it is time to see if something of interest can be done with it. Consider some function of time $f(t)$ that might very well be the voltage signal at the output of a microphone preamplifier. What does one obtain by finding the area under the curve that represents the product of $f(t)$ with $\delta(t - t_0)$? In answering this question, we first make a mathematical statement of the problem with y representing the answer that is sought.

$$y = \int_{-\infty}^{+\infty} f(t)\delta(t - t_0)dt \qquad (22\text{-}65)$$

Now $f(t)$ is a physically generated signal. As such, $f(t)$ is always real, finite, and does not have any exotic mathematical behavior. The chosen unit impulse, however, is zero everywhere except at $t = t_0$. When $t = t_0$, $f(t)$ has the value $f(t_0)$ with $f(t_0)$ being the output voltage of the microphone preamplifier at the instant in time equal to t_0. Mathematically, the voltage at the instant t_0 is just a constant so Eq. 22-65 can be written in the form

$$\begin{aligned} y &= \int_{-\infty}^{+\infty} f(t)\delta(t - t_0)dt \\ &= \int_{-\infty}^{+\infty} f(t_0)\delta(t - t_0)dt \\ &= f(t_0)\int_{-\infty}^{+\infty} \delta(t - t_0)dt \\ &= f(t_0) \end{aligned} \qquad (22\text{-}66)$$

In words, the area under the product curve of $f(t)$ with a unit impulse located at the instant of time t_0 is just the value of the given function of time at the particular instant t_0. The application of the unit impulse in this manner to the continuously time varying signal at the preamplifier output has sifted out or sampled the particular value that exists at the instant in time when t is equal to t_0. Suppose now one desires to sample the microphone preamplifier's output periodically with a sampling period T_s or a sampling frequency $f_s = 1/T_s$. Analytically, this can be stated as

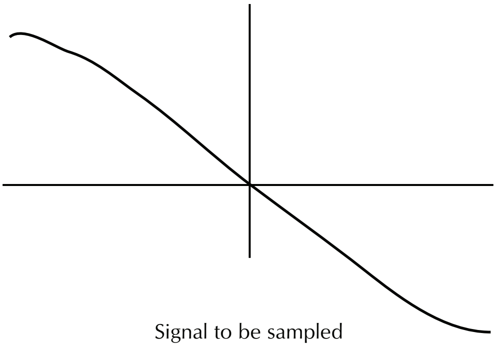
$$\begin{aligned} y(nT_s) &= \sum_{n=-\infty}^{n=+\infty} \int_{-\infty}^{+\infty} f(t)\delta(t - nT_s)dt \\ &= \sum_{n=-\infty}^{n=+\infty} f(nT_s) \end{aligned} \qquad (22\text{-}67)$$

where,

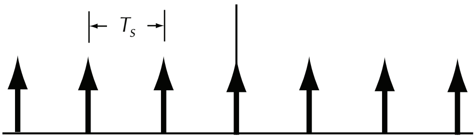
n is any integer.

This process is called periodic impulse sampling. The results of Eq. 22-67 can also be presented graphically at least for a finite number of samples. This is done in Fig. 22-15 wherein are displayed three graphs representing the signal to be sampled, the sampling impulse train, and the sampled results, respectively. The sampled points are exaggerated for ease in viewing.

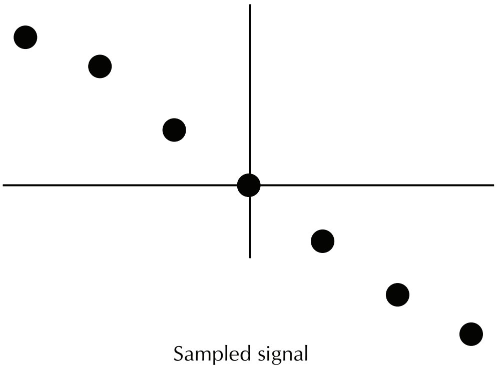
In viewing Fig. 22-15 it should be noted that the time coordinate of each sample matches that of the corresponding sampling impulse while the ordinate or value of the sample is that of the continuous signal at the instant of sampling.



Signal to be sampled



Train of impulses



Sampled signal

Figure 22-15. An example of impulse sampling.

22.1.6 Spectrum of a Periodic Impulse Sampled Signal

What we refer to as the spectrum of a time dependent signal is the description of that signal in terms of frequency rather than in terms of time. One obtains the spectrum of a time dependent signal by calculating its Fourier transform through the application of Eq. 22-53.

On the other hand, if one has on hand a detailed knowledge of a signal's spectrum then it is possible to calculate the time dependence of the signal by calculating the inverse Fourier transform as prescribed in Eq. 22-54. The time signal $f(t)$ and its spectrum $F(\omega)$ constitute a Fourier transform-inverse Fourier transform pair and are compactly denoted as

$$f(t) \Leftrightarrow F(\omega) \tag{22-68}$$

In order to determine the spectrum of a periodic impulse sampled signal we will require just a few more tools. Firstly, let us calculate the Fourier transform or spectrum of a unit impulse. An application of Eq. 22-53 to a unit impulse in the time domain yields

$$\begin{aligned}\Delta(\omega) &= \int_{-\infty}^{\infty} \delta(t)e^{-j\omega t}dt \\ &= e^{-j\omega 0}\int_{-\infty}^{\infty} \delta(t)dt \\ &= 1\int_{-\infty}^{\infty} \delta(t)dt \\ &= 1\end{aligned} \tag{22-69}$$

In Eq. 22-69, $\Delta(\omega)$ represents the Fourier transform of $\delta(t)$ and the second line is justified because the integrand is zero everywhere except at the origin where the exponential term is a constant equal to one. The result is that the spectrum of the unit impulse is a constant independent of ω. This means that all frequencies appear in this spectrum to exactly the same degree. The spectrum is flat. Furthermore, as $\Delta(\omega)$ is purely real, all frequency components have zero phase. This is an example of a limiting case expressed in the uncertainty principle of Eq. 22-59 where the duration in time has tended to zero forcing the bandwidth to become infinite. In conclusion then,

$$\delta(t) \Leftrightarrow 1 \tag{22-70}$$

Suppose, however, that the unit impulse is located at t_0 rather than at the origin. The result in this case can be obtained by direct calculation of course but it is not necessary to do so. One can make use of the general time shift property of Fourier transforms. A time shift to the right of an amount t_0

in the time domain simply multiplies the original transform by $e^{-j\omega t_0}$. This being the case,

$$\delta(t-t_0) \Leftrightarrow 1e^{-j\omega t_0} \tag{22-71}$$

This spectrum is still flat as the magnitude is still unity but now there is a frequency dependent phase $\varphi = -\omega t_0$. Thus the spectrum is now complex. This is an example of another general property of Fourier transforms in that the impulse being located only on the positive time axis means that the time function is neither even nor odd leading to a transform that is complex.

Unit impulses may be multiplied by constants having both numerical value as well as dimensions. In such an instance the impulse is said to have a strength equal to the numerical value of the constant and the product acquires the dimensions of the constant as well. For example, if the constant is k, then Eq. 22-70 and Eq. 22-71 would become

$$\begin{aligned} k\delta(t) &\Leftrightarrow k \\ k\delta(t-t_0) &\Leftrightarrow ke^{-j\omega t_0} \end{aligned} \tag{22-72}$$

The concept of impulse can be carried over to the frequency domain as well. In this application, the independent variable is angular frequency rather than time while the mathematical properties of the unit impulse are unchanged. An impulse located at the origin of coordinates is symbolized as $\delta(\omega)$. An impulse located at a fixed angular frequency ω_0 would be written as $\delta(\omega - \omega_0)$. An infinite sequence of impulses with a constant separation could be denoted as

$$\sum_{n=-\infty}^{n=+\infty} \delta(\omega - n\omega_0)$$

where,

n is any integer.

These impulses are described graphically in Fig. 22-16.

We are now in position to calculate the expressions in the time domain that correspond to the above spectra in the frequency domain. This is done by applying in turn Eq. 22-54 to each of the above spectra.

For the impulse at the origin

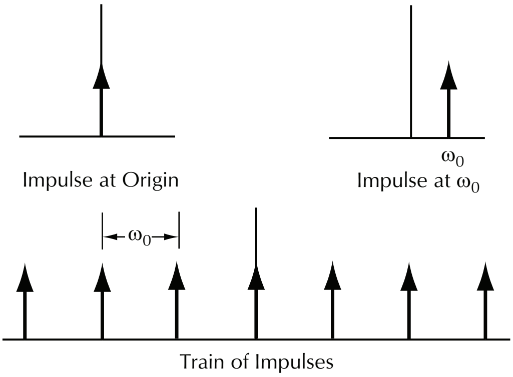


Figure 22-16. Impulses in the frequency domain.

$$\begin{aligned} f(t) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \delta(\omega)e^{j0t}d\omega \\ &= \frac{1}{2\pi}e^{j0t} \\ &= \frac{1}{2\pi} \end{aligned} \tag{22-73}$$

For the displaced impulse at ω_0

$$\begin{aligned} f(t) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \delta(\omega-\omega_0)e^{j\omega t}d\omega \\ &= \frac{1}{2\pi}e^{j\omega_0 t} \end{aligned} \tag{22-74}$$

Finally, for the infinite train of impulses

$$\begin{aligned} f(t) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \sum_{n=-\infty}^{n=+\infty} \delta(\omega-n\omega_0)e^{j\omega t}d\omega \\ &= \frac{1}{2\pi}\sum_{n=-\infty}^{n=+\infty} e^{jn\omega_0 t} \end{aligned} \tag{22-75}$$

In summary, the Fourier transform-inverse Fourier transform pairs are respectively

$$\frac{1}{2\pi} \Leftrightarrow \delta(\omega) \tag{22-76}$$

$$\frac{1}{2\pi}e^{j\omega_0 t} \Leftrightarrow \delta(\omega-\omega_0) \tag{22-77}$$

$$\frac{1}{2\pi}\sum_{n=-\infty}^{n=+\infty} e^{jn\omega_0 t} \Leftrightarrow \sum_{n=-\infty}^{n=+\infty} \delta(\omega-n\omega_0) \tag{22-78}$$

The sampling signal $s(t)$ of Eq. 22-64 is a periodic pulse train of period T_s with a fundamental frequency $f_s = 1/T_s$ and an attendant angular frequency $\omega_s = 2\pi f_s$. This being the case, $s(t)$ can also be written in the form of the Fourier exponential series of Eq. 22-37 with ω_s replacing ω_0.

$$\begin{aligned} s(t) &= \sum_{n=-\infty}^{n=+\infty} \delta(t - nT_s) \\ &= \sum_{n=-\infty}^{n=+\infty} c_n e^{jn\omega_s t} \end{aligned} \tag{22-79}$$

where,

$$c_n = \frac{1}{T_s} \int_{-\frac{T_s}{2}}^{+\frac{T_s}{2}} s(t) e^{-jn\omega_s t} dt \tag{22-80}$$

In writing Eq. 22-80 we have made use of the fact that the integration need be carried out only over one period of the periodic signal. Our periodic sampling signal $s(t)$ is an even function of time so we have chosen the interval of integration to straddle the origin. This choice makes the coefficient calculation quite simple. The sampling signal in this time interval need be represented by only the unit impulse at the origin hence, for all n Eq. 22-80 becomes

$$\begin{aligned} c_n &= \frac{1}{T_s} \int_{-\frac{T_s}{2}}^{+\frac{T_s}{2}} \delta(t) e^{-jn\omega_s t} dt \\ &= \frac{1}{T_s} \int_{-\frac{T_s}{2}}^{+\frac{T_s}{2}} \delta(t) e^{-jn\omega_s 0} dt \\ &= \frac{1}{T_s} \end{aligned} \tag{22-81}$$

As a consequence, we now have two equivalent forms for representing $s(t)$ as expressed in

$$\begin{aligned} s(t) &= \sum_{n=-\infty}^{n=+\infty} \delta(t - nT_s) \\ &= \frac{1}{T_s} \sum_{n=-\infty}^{n=+\infty} e^{jn\omega_s t} \end{aligned} \tag{22-82}$$

We are now in position to determine the spectrum or Fourier transform of the sampling signal. We will do so by first noting that because of linearity the transform pair in Eq. 22-78 can be scaled by any constant factor. Furthermore ω_0 can be identified as ω_s as they are simply constants. A rescaling of the pair in Eq. 22-78 by a multiplying factor $2\pi/T_s$ and replacing ω_0 by ω_s results in

$$\frac{1}{T_s} \sum_{n=-\infty}^{n=+\infty} e^{jn\omega_s t} \Leftrightarrow \frac{2\pi}{T_s} \sum_{n=-\infty}^{n=+\infty} \delta(\omega - n\omega_s) \tag{22-83}$$

Now note that on the left side of the arrow in Eq. 22-83 is just the sampling signal itself while the expression on the right side is its Fourier transform or spectrum. Stated in formal terms

$$\begin{aligned} S(\omega) &= \int_{-\infty}^{+\infty} s(t) e^{-j\omega t} dt \\ &= \frac{2\pi}{T_s} \sum_{n=-\infty}^{n=+\infty} \delta(\omega - n\omega_s) \end{aligned} \tag{22-84}$$

Finally, Eq. 22-83 could equally as well be written as

$$\sum_{n=-\infty}^{n=+\infty} \delta(t - nT_s) \Leftrightarrow \frac{2\pi}{T_s} \sum_{n=-\infty}^{n=+\infty} \delta(\omega - n\omega_s) \tag{22-85}$$

In words, the spectrum of an infinite sequence of uniformly spaced unit impulses in the time domain is an infinite sequence of uniformly spaced impulses of strength $2\pi/T_s$ in the frequency domain. The spacing in the two domains is related through T_s being equal to $2\pi/\omega_s$. Fig. 22-17 illustrates the spectrum of the sampling signal.

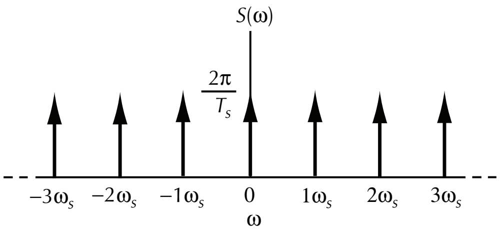


Figure 22-17. The spectrum $S(\omega)$ of the sampling signal $s(t)$.

We require just one more tool in order to accomplish the goal of this section. Reference was made to this tool earlier while discussing the process of convolution. This tool is known as the Frequency Convolution Theorem of Fourier transforms. This theorem states that given two functions of time, say $f(t)$ and $g(t)$, with the Fourier transform of $f(t)$ being $F(\omega)$ and that of $g(t)$ being $G(\omega)$, then the Fourier transform of the product of $f(t)$ with $g(t)$ is given by the convolution of $F(\omega)$ with $G(\omega)$ divided by 2π. Expressed in the concise language of mathematics,

$$f(t)g(t) \Leftrightarrow \frac{1}{2\pi}F(\omega) * G(\omega) \tag{22-86}$$

For the present case of interest, which involves sampling the signal at the output of the microphone preamplifier, the product at the left of Eq. 22-86 is $f(t)$ with $s(t)$. In this instance we are taking $f(t)$ to be the continuous time description of the microphone signal of interest and $s(t)$ to be the periodic impulse train of Eq. 22-82. Following the pattern of Eq. 22-86 then, the Fourier transform or spectrum of this product, which we denote as $F_s(\omega)$, will be given by

$$\begin{aligned} F_s(\omega) &= \frac{1}{2\pi}F(\omega) * S(\omega) \\ &= \frac{1}{2\pi}F(\omega) * \frac{2\pi}{T_s}\sum_{n=-\infty}^{n=+\infty} \delta(\omega - n\omega_s) \end{aligned} \tag{22-87}$$

$F(\omega)$ is of course the Fourier transform of $f(t)$ and is unknown unless one specifies what particular sounds the microphone is being exposed to. Rather than look at the results for a specific case, it is much more meaningful to consider a generic microphone signal spectrum that might serve as a bounding one for audio signals. Audible audio signals can and do extend to quite low frequencies so to be on the safe side one might consider the lower limit of the spectrum to be zero. Furthermore, the most energetic signals occur in the range of 500 Hz and below while the strength above 500 Hz usually diminishes as the frequency increases and is practically zero beyond 20 kHz. The audible audio spectrum is thus naturally band limited even when artificial limits are not imposed. This being the case, a generic $F(\omega)$ might appear as in Fig. 22-18 where ω_m represents the maximum angular frequency at the upper limit.

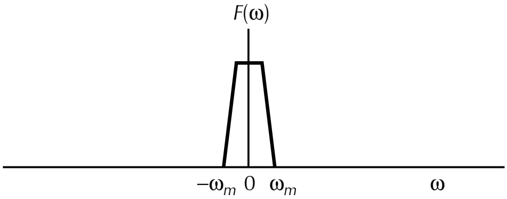


Figure 22-18. Spectrum of a generic audible audio signal.

With $F(\omega)$ now at hand, it is at last possible to determine the spectrum of the sampled microphone signal by invoking Eq. 22-87. Remember that in the convolution process one basically sweeps one function over the other while at each step in the process one calculates the area under the product curve in the region of overlap of the two functions. Graphically then we are sliding the train of impulse functions of Fig. 22-17 across the spectrum of Fig. 22-18. At each of the points of overlap, the impulses simply replicate and scale $F(\omega)$, producing the final result displayed in Fig. 22-19.

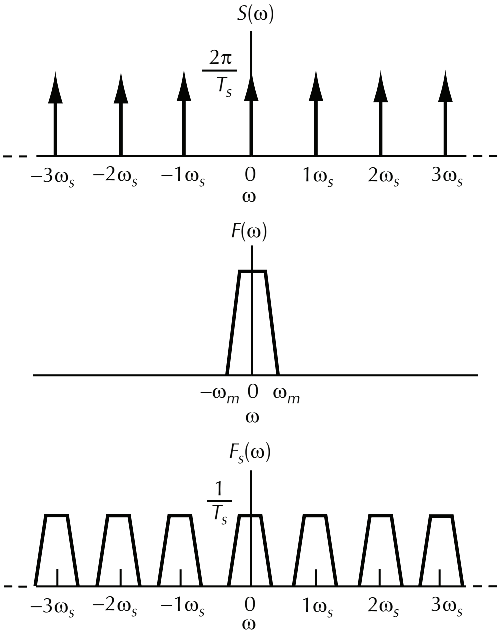
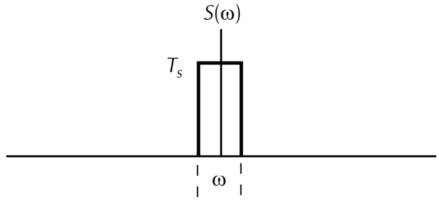
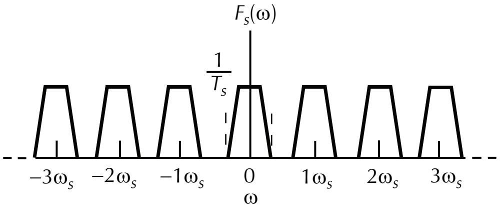


Figure 22-19. Spectra involved in the convolution as well as the result.

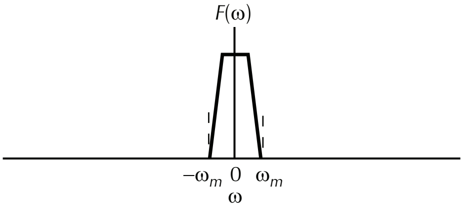
The spectrum of the impulse sampled signal, $F_s(\omega)$, is probably a surprising result in that it contains not only a scaled version of the spectrum of the continuous microphone signal but also an infinite number of replications uniformly spaced at integral multiples of the sampling angular frequency. At this point it would be reasonable to consider what step or steps must be taken to extract from $F_s(\omega)$ just the spectrum of the original microphone signal itself, that is, $F(\omega)$. The answer is supplied in Fig. 22-20 where an ideal low pass filter with a cutoff angular frequency slightly greater than ω_m and a scaling factor of T_s is applied to the spectrum $F_s(\omega)$. The spectrum so filtered is just $F(\omega)$.


$S(\omega)$
T_s
ω

A. Low Pass Filter Response.


$F_s(\omega)$
$\frac{1}{T_s}$
$-3\omega_s$
$-2\omega_s$
$-1\omega_s$
0
$1\omega_s$
$2\omega_s$
$3\omega_s$
ω

B. Spectrum to be Filtered.


$F(\omega)$
$-\omega_m$
0
ω_m
ω

C. Filtered Result.

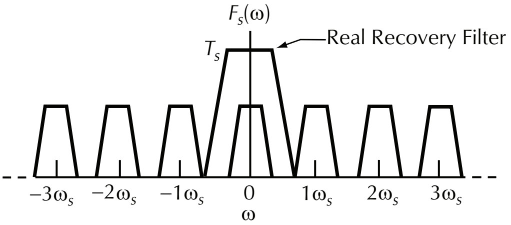
Figure 22-20. Recovery of original spectrum.

Fig. 22-20A depicts the response of the required low pass filter. Remember that the Fourier description of a low pass filter involves both positive and negative frequencies and hence the filter is symmetrical about the origin. Fig. 22-20B is the spectrum of the impulse sampled signal while Fig. 22-20C is the result of the filtering process and is the same as the spectrum of the original microphone signal. This spectrum contains all of the information of the original continuous time dependent signal $f(t)$ and nothing has been lost as a result of the sampling process when carried out as prescribed above.

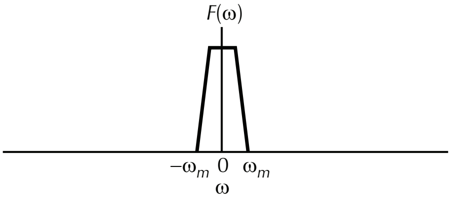
22.1.7 Sampling Theorem

A visual inspection of Fig. 22-20 should indicate that the sampling angular frequency ω_s is considerably more than twice the maximum angular frequency in the program material ω_m. The spacing is sufficiently large, in fact, that an ideal filter is not required to perform the recovery. This fact is illustrated in Fig. 22-21.

Fig. 22-21A depicts the sampled spectrum with the response of a physically realizable or real filter superimposed about the portion of the spectrum to be recovered while Fig. 22-21B illustrates the results of the recovery. The cogent property of a real filter illustrated here is that such a filter possesses a finite attenuation rate or slope at the pass band edge. An ideal filter, of course, would have an infinite slope at the pass band edge.


$F_s(\omega)$
T_s
Real Recovery Filter
$-3\omega_s$
$-2\omega_s$
$-1\omega_s$
0
$1\omega_s$
$2\omega_s$
$3\omega_s$
ω

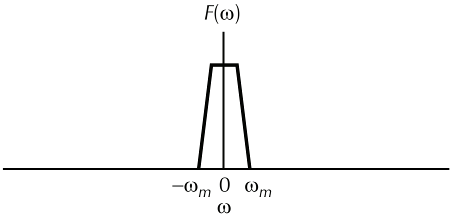
A. Spectrum to be Filtered.


$F(\omega)$
$-\omega_m$
0
ω_m
ω

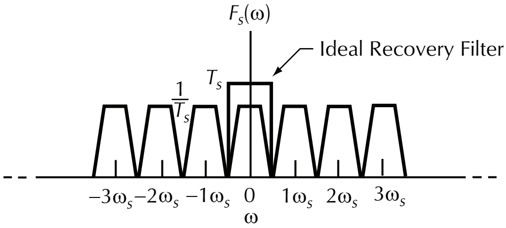
B. Filtered Result.

Figure 22-21. Recovery with a real filter is possible when ω_s is sufficiently large.

One might very well reason after viewing Fig. 22-20 and Fig. 22-21 that with the aid of an ideal filter, operation would be possible with $\omega_s = 2\omega_m$. This situation is presented in Fig. 22-22.


$F(\omega)$
$-\omega_m$
0
ω_m
ω

A. Original Spectrum.


$F_s(\omega)$
Ideal Recovery Filter
T_s
$\frac{1}{T_s}$
$-3\omega_s$
$-2\omega_s$
$-1\omega_s$
0
$1\omega_s$
$2\omega_s$
$3\omega_s$
ω

B. Sampled Spectrum.

Figure 22-22. Spectra appearance when $\omega_s = 2\omega_m$.

Upon reviewing Fig. 22-22 the first guess might be that the ideal filter would allow recovery of just the central spectrum alone and successful operation would be possible when the sampling frequency is just twice the maximum frequency in the original time dependent signal of interest. This, however, is not the case as is evident from the following illustration taken from the time domain.

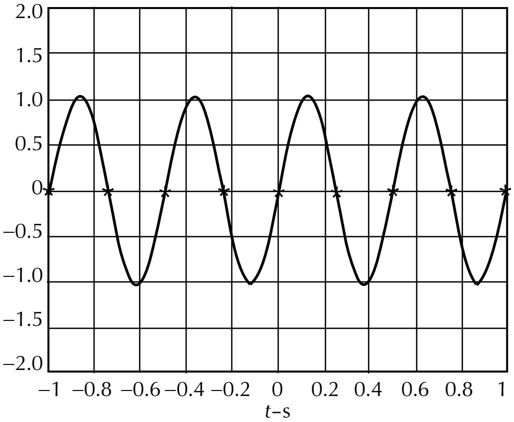


Figure 22-23. Time domain when $\omega_s = 2\omega_m$.

In Fig. 22-23 the signal is a sinusoid with a frequency of 20 kHz. The sampling frequency itself is set at precisely 40 kHz so that two samples are taken during each period of the signal. In the situation depicted in the figure, the sampling signal is in phase with the signal to be sampled so that the samples are taken at the zero crossings of the signal of interest. In this instance, then, all of the samples are zero and nothing of value is learned about the signal. Clearly, such a sampling process could never lead to the recovery of the original signal from the sampled values.

After examination of the foregoing examples we are now in position to state the Sampling Theorem. In simple terms the Sampling Theorem states that if a continuous time signal is band limited to a maximum frequency f_m, then the continuous time signal can be uniquely determined from a uniformly spaced sequence of samples taken at a rate greater than twice f_m. In short, $f_s > 2f_m$. The corresponding angular frequencies are of course just 2π times these values.

22.1.8 Aliasing

Aliasing is a phenomenon that occurs whenever the sampling frequency is less than twice the maximum frequency in the sampled spectrum. That is, $f_s < 2f_m$. Here we are using the actual frequencies rather than the angular frequencies for the sake of clarity. Most of us first experience this phenomenon when viewing western movies that feature stagecoaches or wagons in motion. Conventional 35 mm sound movies are made up of sampled data sequences consisting of individual pictures or frames made at the rate of 24 frames per second (fps). As long as the spoked wheels are rotating slowly so that the spokes pass a reference point at a rate that is less than one-half of the frame rate, the spoke motion appears to be consistent with the linear translation of the vehicle. Once the spoke rate exceeds one-half of the frame rate, the spoke motion appears to be too slow as compared with the linear translation. In this instance, then, the spokes appear to be advancing at a lower rate or frequency than required to account for the vehicle motion.

In fact, when the spoke rate is first equal to the frame rate, the spokes appear to be standing still while the vehicle is moving at a good clip. As if this is not bad enough, when the spoke rate first exceeds the frame rate, the rotation of the spokes appears to be in the opposite direction than that required to be consistent with the vehicle motion! As a small child and a big western movie fan, this bothered me no end as I saw moving spoked vehicles every day in real life and this phenomenon did not occur. This same phenomenon occurs while observing rotating machinery with the aid of a strobe light. In this instance, however, the strobe frequency is usually adjusted to make the rotational motion to appear to stand still or move at a slow rate.

Aliases are frequency components not present in the original signal that are generated whenever the sampling frequency is less than twice the maximum frequency that appears in the original signal. Fig. 22-24 displays a portion of a sampled spectrum where this is the case.

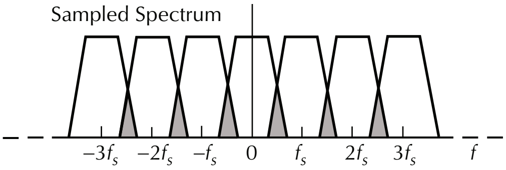


Figure 22-24. Aliasing occurs in the shaded regions.

The situation depicted in Fig. 22-24 results when the sampling frequency is too low. In this instance the replica spectra overlap their nearest neighbors as indicated by the shaded regions in the figure. The regions in which aliasing occur are given by

$$f = nf_s \pm f_m \tag{22-88}$$

where,

n is ±1, ±2, ±3, etc.

The portion of the spectrum that spans between $\pm f_m$ is the region occupied by the original signal. Aliases occurring anywhere in this region prevent the recovery of an undistorted version of the original signal through simple filtering techniques. Aliasing details are best presented through a numerical example.

Fig. 22-25 displays only the positive frequency axis because of space considerations. The situation on the negative frequency axis is just the mirror image of what is displayed in the figure. Additionally, there is room in the figure for only one replica spectrum centered about the sampling frequency of 25 kHz. The original spectrum which is often referred to as the baseband spectrum in communication work is centered on the origin and extends from –20 kHz to +20 kHz with only the positive portion appearing in the figure. Aliasing occurs in the region between 5 kHz and 20 kHz. The sloping line indicated by the lone arrow represents information associated with the high frequency region of the replica spectrum, in fact that region between 12.5 kHz and 20 kHz. Note, however, that it falls between 12.5 kHz and 5 kHz in the baseband spectrum. Specifically, 20 kHz in the replica appears as 5 kHz in the baseband. In other words 20 kHz is aliased or masquerades as 5 kHz and would appear as 5 kHz in the recovery process. Similarly, 19 kHz would appear as 6 kHz, 18 kHz would appear as 7 kHz, etc. until the foldover frequency of $f_s/2$ or 12.5 kHz is reached.

Beyond the foldover frequency, the progression just reverses until 20 kHz is reached. Any attempt to recapture the entire baseband spectrum will involve a filter that extends to 20 kHz or slightly beyond and hence will include the alias region. If one compares such a recovered spectrum with the original, it will be found that the shape of the spectrum will be highly distorted from 5 kHz and beyond. All of this is avoided by satisfying the conditions set forth in the Sampling Theorem. If one is at liberty to select the sampling frequency, one simply samples at a rate greater than twice the value of the maximum frequency in the program material.

The answer to the question of how much greater hinges on the sharpness of the available recovery filter with low order recovery filters requiring higher sampling frequencies. The situation might be, however, that the sampling frequency is beyond one's control. In this instance the program material to be sampled must be band limited by what is called an anti-alias filter to a frequency maximum less than one-half of the available sampling frequency. Here the answer to the question of how much less depends on the sharpness of the available anti-alias filter. Anti-alias filters are almost universally employed whether needed in principle or not to guard against possible high frequency noise pollution of the program material to be sampled.

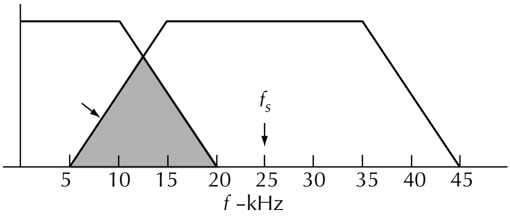


Figure 22-25. Aliasing details.

22.1.9 Realistic Sampling

Man-made sampling pulses and techniques fall short of the theoretical ideal of the unit impulse. Fortunately, this does not invalidate the Sampling Theorem or the general conclusions that stem from the Sampling Theorem. When one employs an accurate mathematical description of the actual man-made periodic sampling pulses employed the results and conclusions follow the general outline of the ideal case presented in the foregoing development of sampling. There are some differences of course in the mathematical details, none of which are debilitating. As an example of a realistic sampling technique we will examine the details associated with a process known as sample and hold. This is an often employed technique particularly where it is desired to convert the analog samples into a quantized format that is subsequently encoded into a digital word through a process known as pulse code modulation or PCM. Fig. 22-26 is an example of a typical sample and hold circuit.

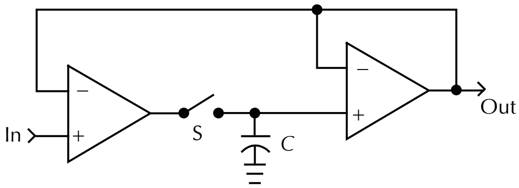


Figure 22-26. Sample and hold circuit.

The switch S in the circuit of Fig. 22-26 is actually a logic operated electronic switch that is briefly closed and then opened at the beginning of the sampling period. While the switch is closed the capacitor is charged to a voltage equal to the signal voltage. The capacitor holds this voltage after the switch opens for the remainder of the sampling period while the buffer amplifier supplies this same voltage to subsequent circuitry and isolates the hold capacitor. The waveform at the output is of the form of a staircase with the elevations of the steps corresponding to the values of the signal voltage at the respective instants of sampling. The time domain behavior of this process is displayed in Fig. 22-27.

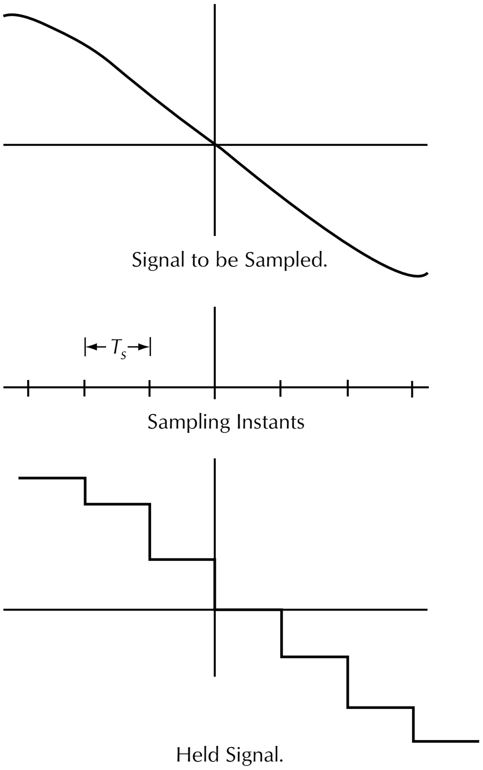


Figure 22-27. Time domain behavior of sample and hold.

From a mathematical standpoint, the staircase time domain signal of Fig. 22-27 is given by the convolution of a unit amplitude rectangular pulse of width T_s with the periodic impulse sampled signal of Eq. 22-67. If we denote the pulse waveform as $p(t)$, then $p(t)$ is described by

$$\begin{aligned} p(t) &= 1, \text{ when } 0 < t < T_s \\ p(t) &= 0, \text{ when } t > T_s \end{aligned} \qquad (22\text{-}89)$$

Formally, the staircase time domain behavior denoted as $f_h(t)$ is given by

$$f_h(t) = p(t) * \sum_{n=-\infty}^{n=+\infty} f(nT_s) \qquad (22\text{-}90)$$

The quantity of interest however is not $f_h(t)$ but rather its Fourier transform $F_h(\omega)$. $F_h(\omega)$ is the spectrum of the sample and hold signal. Knowledge of this spectrum is needed in detail in order to formulate techniques for recovering the spectrum of the original unsampled signal. From Eq. 22-90, $f_h(t)$ is given by the convolution of two functions of time. Recall that the Fourier transform of the convolution of two time functions is simply the product of the transforms of the individual time functions. $F_h(\omega)$ can be calculated by taking the product of the Fourier transform of $p(t)$ with the Fourier transform of the infinite sum of Eq. 22-90. This latter transform is just the spectrum $F_s(\omega)$ as displayed in Fig. 22-19 and expressed in Eq. 22-87.

$$F_s(\omega) = \frac{1}{T_s} \sum_{n=-\infty}^{n=+\infty} F(\omega - n\omega_s) \qquad (22\text{-}91)$$

The calculation of the Fourier transform of $p(t)$ is quite straightforward.

$$\begin{aligned} P(\omega) &= \int_{-\infty}^{+\infty} p(t) e^{-j\omega t} dt \\ &= \int_{0}^{T_s} 1 e^{-j\omega t} dt \\ &= \frac{1 - e^{-j\omega T_s}}{j\omega} \end{aligned} \qquad (22\text{-}92)$$

Now, if one factors the term $e^{-j\omega(T_s/2)}$ from the last expression of Eq. 22-92 and employs the exponential identity for the sine function, then Eq. 22-92 becomes

$$P(\omega) = T_s \left| \frac{\sin\left(\omega \frac{T_s}{2}\right)}{\omega \frac{T_s}{2}} \right| e^{-j\omega \frac{T_s}{2}} \qquad (22\text{-}93)$$

It should be apparent from Eq. 22-93 that the spectrum of the unit amplitude pulse is complex. This pulse is a real function existing only on the positive time axis and hence is not an even or an odd function. For such functions, the Fourier transform is complex having both a magnitude as well as a phase. The phase behavior is linear and describes a pure time shift along the positive t axis of an amount $T_s/2$. This in effect accounts for the fact that the pulse does not symmetrically straddle the origin but rather begins immediately to the right of the origin. The magnitude of the spectrum is the pulse duration multiplied by the absolute value of a sinc function, that is a function of the form $\sin(x)/x$. With both $P(\omega)$ and $F_s(\omega)$ at hand we can at last write an expression for the spectrum of the sample and hold signal.

$$F_h(\omega) = T_s \left| \frac{\sin\left(\omega \frac{T_s}{2}\right)}{\omega \frac{T_s}{2}} \right| e^{-j\omega \frac{T_s}{2}} \frac{1}{T_s} \sum_{n=-\infty}^{n=+\infty} F(\omega - n\omega_s) \tag{22-94}$$

Fig. 22-28 displays $F_s(\omega)$ as overlaid by the magnitude of $P(\omega)$. The product of these two curves describes $F_h(\omega)$.

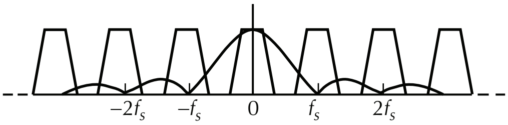


Figure 22-28. The sample and hold spectrum is the product of the two curves.

The two components displayed in Fig. 22-28 consist of the spectrum that would be produced by a pure unit impulse sampling of our original audio signal and the spectrum associated with the unit amplitude hold pulse. The spectrum of our original audio signal resulting from sampling by means of the sample and hold process is given by the point-by-point product of the two component curves. The magnitude of the sinc function beneficially attenuates the repeated replicas but at the same time also alters the shape of the baseband spectrum.

A viable low pass filter intended for baseband spectrum recovery must accomplish two things. It must remove or seriously attenuate the replica spectra that appear at multiples of the sampling frequency and it must compensate for the attenuation introduced by the sinc function throughout the baseband region. This means that the recovery filter must have just the inverse response of the sinc function in the frequency region between minus and plus the maximum frequency of the original audio signal. Beyond the frequency range of the baseband spectrum the recovery filter must further remove the remaining remnants of the replica spectra.

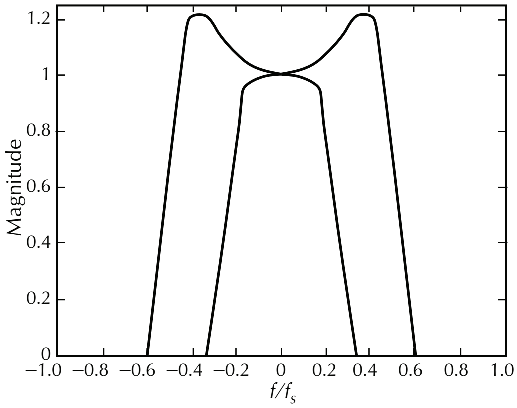


Figure 22-29. Recovery filter response and baseband spectrum after sample and hold.

Fig. 22-29 displays the distorted shape of the baseband spectrum resulting from the hold operation. The recovery filter corrects this distortion. The response of the recovery filter is tailored to match the inverse of the sinc function spectrum over the range between $\pm f_m$. Beyond the baseband frequency range the recovery filter furnishes simple attenuation. In the example at hand, f_m is one-third of the sampling frequency f_s. It should be noted that in this case as well as in the previous examples the drawings have been made with linear rather than with logarithmic axes. In all instances some license has been taken in making the drawings in that the spectra do not attenuate to zero but rather to some negligibly small value.

22.2 Analog to Digital Conversion

In the foregoing discussion the emphasis has been on the spectral characteristics of the sampling process and the steps that must be taken in order to recover the original continuous time dependent signal from the sampled sequence. There is no practical reason to perform sampling followed by immediate recovery other than perhaps as an exercise. The collection of a sampled data sequence is a precursor to encoding the information so gathered into a form that can be recognized and processed by computer circuitry. From a mathematical point of view computer circuitry is based upon the two valued logic system and algebra of George Boole (1815-1864). The number system that lends itself naturally to computer operations is the binary or modulus two system. Analog audio signals vary continuously between a negative minimum and positive maximum with the difference between these two values constituting a peak to peak value or range. One of the first considerations that must be made in the encoding process is the designation of the least change in the sampled value that is to be recognized by the encoding scheme. This least change sets the numerical precision or quantization of the encoded signal. Precision in a quantity expressed with any modulus hinges on the range, the modulus, and the number of bits involved. Denoting the precision or quantization by Δ, the range by V_{pp}, the modulus by M, and the number of bits by n, the relationship is

$$\Delta = \frac{V_{pp}}{M^n} \tag{22-95}$$

An explanation of Eq. 22-95 and a brief review of the formalism of expressing quantities in any number base are perhaps worthwhile. The concept of positional notation is presented in Table 22-1.

Table 22-1. Positional Notation Table

M^3	M^2	M^1	M^0	·	M^{-1}	M^{-2}	M^{-3}	M^{-4}
–	–	–	–		–	–	–	–

The table is divided into two parts by the radix point. The region to the left of the radix point represents whole numbers while that to the right represents fractional numbers. In the most familiar decimal system $M = 10$ and the positions to the left of what is now the decimal point have weighting factors that are progressively $10^0 = 1$, $10^1 = 10$, $10^2 = 100$, and $10^3 = 1000$. To the right the weighting factors are progressively 1/10, 1/100, 1/1000, and 1/10,000. If then the entries in the second row of the table were 1234.4321, the quantity so represented actually means 1 × 1000 plus 2 × 100 plus 3 × 10 plus 4 × 1 plus 4/10 plus 3/100 plus 2/1000 plus 1/10,000. In the decimal system the column positions are called digit positions with the most significant digit being on the extreme left. In general, the numeric entry in any column can take on any value from zero through $M - 1$. Now in the binary or base two system, $M = 2$ and the entries in any column can only be 0 or 1. The column positions are now called the bit positions with the most significant bit again being on the extreme left. If the entries in the second row of the table were 1010.0101, the quantity represented would be 1 × 8 plus 0 × 4 plus 1 × 2 plus 0 × 1 plus 0/2 plus 1/4 plus 0/8 plus 1/16. The rule governing the structure of the position table is that for any given position the weighting factor is M times as great as that of the position to the immediate right of the given position. Eq. 22-95 is structured so that the position rule is always satisfied when employing a limited numeric code to cover a given range of values. In such an instance the quantization or Δ is the value of the smallest increment associated with the code. In the case of binary codes Δ is the value represented by the least significant bit. Let us examine a simple example. Suppose that we encode a voltage range from 0 to 1.6 V by employing a four bit binary code. In this instance Δ will be 1.6 V divided by 2^4 or 1.6/16 = 0.1 V. The coded values might reasonably be those of Table 22-2.

Table 22-2. Binary Codes Versus Volts

Binary Code	Volts
0000	0
0001	0.1
0010	0.2
0011	0.3
0100	0.4
0101	0.5
0110	0.6
0111	0.7
1000	0.8
1001	0.9
1010	1.0
1011	1.1
1100	1.2
1101	1.3
1110	1.4
1111	1.5

Several observations may be made from examining the table. The step size is indeed 0.1 V and the positional rule is indeed satisfied in that the code 1000 represents eight increments, 0100 represents four increments, 0010 represents two increments, and 0001 represents one increment. A four bit binary code possesses 16 unique combinations of 1s and 0s and hence has 16 levels including the zero level. In general, however, there are only $2^n - 1$ steps, in this case 15, so that the largest value represented by such a code is always the top of the range less one least significant bit. The step size here is coarse because of the small number of bits. A sixteen bit code would have produced a step size of about 24 microvolts (μV) for the same voltage range.

Audio signals of an electrical nature are electric analogs of acoustic pressure and as such are intrinsically bipolar. The coding scheme selected for portraying such signals must retain both algebraic sign and magnitude behavior. At first glance one might consider dedicating one bit to indicate the algebraic sign with the remaining bits being devoted to magnitude. When this is done the most significant bit is usually employed as the sign bit. Such a scheme is viable but does have a serious shortcoming as we shall see.

The most popular signed code employed by microprocessors in performing arithmetic operations is called two's complement binary. The two's complement of a given binary number is obtained by first taking the one's complement of the given binary number and then adding one to it. The one's complement of a binary number is obtained by complementing each bit of the binary number.

Suppose the original decimal number is +7. The normal binary statement with a four bit code would be 0111. The one's complement is obtained by changing each zero to a one and each one to a zero to obtain 1000. The two's complement is obtained from this intermediate step simply by adding one to it so that 1000 becomes 1001.

In two's complement coding positive numbers are described by straight binary while negative numbers are described by the twos complement of the corresponding positive number. As a consequence the most significant bit is again the sign bit with a zero representing positive numbers while a one represents negative numbers. Finally another scheme called offset binary is similar to two's complement except that the most significant bit is complemented with a one representing positive numbers and a zero representing negative numbers. The following table displays each of these schemes for four binary bits covering a voltage range from −0.8 to +0.8 V or 1.6 $V_{\text{peak to peak}}$.

An inspection of the sign and magnitude column of Table 22-3 reveals immediately the shortcoming of this particular coding scheme in that there are two codes representing the value of zero depending upon whether zero is approached from values less than zero or from values greater than zero. This ambiguity is removed by employing either the twos complement code or the offset binary code. It is also of note that one can readily change from the twos complement code to the offset binary code by simply complementing the most significant bit.

Table 22-3. Binary Codes for Bipolar Signals

Decimal Value	Sign and Magnitude	Two's Complement	Offset Binary	Voltage (Volt)
+7	0111	0111	1111	+0.7
+6	0110	0110	1110	+0.6
+5	0101	0101	1101	+0.5
+4	0100	0100	1100	+0.4
+3	0011	0011	1011	+0.3
+2	0010	0010	1010	+0.2
+1	0001	0001	1001	+0.1
+0	0000	0000	1000	0
−0	1000	0000	1000	0
−1	1001	1111	0111	−0.1
−2	1010	1110	0110	−0.2
−3	1011	1101	0101	−0.3
−4	1100	1100	0100	−0.4
−5	1101	1011	0011	−0.5
−6	1110	1010	0010	−0.6
−7	1111	1001	0001	−0.7
−8	None	1000	0000	−0.8

Many different techniques for performing analog to digital conversion have been developed over the years. In audio applications there are two markedly different techniques that have proven their worth. The first of these is the method of successive approximations that held the stage almost exclusively for many years. The second of these is the delta-sigma technique that promises to eventually dominate in audio applications. Both of these converter techniques employ a special circuit known as a digital to amplitude converter or DAC as a component in the overall analog to digital converter structure. Analog to digital converters, regardless of the technique of conversion, are referred to collectively as ADCs. A DAC will be discussed initially as it is a crucial element regardless of the selected ADC technique. Fig. 22-30 is a simple circuit for a four bit DAC.

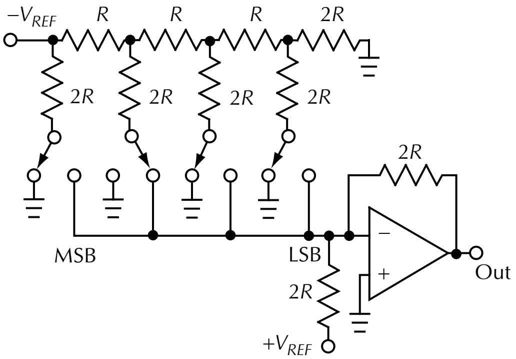


Figure 22-30. Four bit DAC.

The DAC displayed in Fig. 22-30 is configured to operate with an offset binary bipolar code. It will operate equally as well with a twos complement code if the MSB data line includes an inverter. The switches employed are electronic switches activated by the logic levels on the data lines though they have been drawn as conventional SPDT for simplicity. The ground position on each switch means that the respective data bit is a zero while the connection to the opamp summing bus means that the respective data bit is a one. The switches as shown are in response to the code 0100.

The heart of the circuit is the R–$2R$ ladder network. The resistance to ground as measured at the negative voltage reference terminal is R independent of the settings of the switches because the summing bus of the opamp is itself a virtual ground. The total current directed into the negative voltage reference source is the value of the negative reference voltage divided by R. The current in the MSB switch is ½ of the total. The current for the next bit switch is ¼ of the total, the next ⅛, and finally the LSB switch has 1⁄16 of the total current. The remaining 1⁄16 of the total current directed into the negative voltage reference source always exists in the $2R$ resistance that terminates the ladder network. The offset binary code for the value of zero is 1000. The connection of the summing bus to a positive voltage reference equal in magnitude to the negative voltage reference through a resistance forces the opamp output to zero volts. With bipolar reference voltages of +0.8 V and −0.8 V, the DAC output tracks the last column of Table 22-3.

The DAC circuit was drawn for a four bit device as this is sufficient for an understanding of the circuit operation. A simple extension of the R–$2R$ ladder network allows for more bits to be handled in the same manner. Increasing the number of bits employed increases the dynamic range of signals that may be processed by either a DAC or ADC. The dynamic range mentioned here refers to the ratio of the peak signal value to the least signal value that may be handled by the encoding process. This is optimistically stated as being 2^n to 1 where n is the number of bits. For bipolar signals, however, the most significant bit is devoted solely to representing the algebraic sign of the signal and hence only $n - 1$ bits are available for magnitude assignments. This being the case, a more conservative statement of the dynamic range is 2^{n-1} to one. Upon letting D represent the dynamic range in decibels then

$$D = 20\mathrm{dB}\log\left(\frac{2^{n-1}}{1}\right) \tag{22-96}$$

Sixteen bit encoding is an often encountered standard for high quality audio program material. When this is the case Eq. 22-96 yields a dynamic range of 90 dB. As a caution, this dynamic range limitation is strictly associated with the linear encoding process and is a separate issue from the overall dynamic range of a complete system. A linear encoding process has an uncertainty of ±½ LSB that contributes to what is termed quantization noise. This quantization noise acts to worsen a system's signal to noise ratio. This fact alone would place a premium on a large number of bits.

Fig. 22-31 is a block diagram of a complete ADC that employs the technique of successive approximation conversion.

In addition to the elements previously discussed, the successive approximation converter contains a comparator, a special successive approximation register, an output data register, a master clock, and a control unit. The control unit in collaboration with the master clock establishes the sample rate and generates the timing and logic signals necessary for the overall system operation. The comparator compares the voltage value of the hold signal with the output voltage of the DAC and generates a signal only when the DAC voltage exceeds the hold voltage. At the start of a conversion as soon as the sample and hold circuit adopts the hold mode the SAR sets the most significant bit in the DAC via means of the data bus. The most significant bit forces the DAC output to assume a value equal to one-half of the total voltage range. If the hold signal exceeds this value, the comparator does not trip and the SAR continues asserting the most significant bit and then sets the next most significant bit.

The action continues in this fashion until the comparator trips. When this occurs, the last bit set is returned to zero and the next bit down the line is set. The operation continues in this manner until all bits

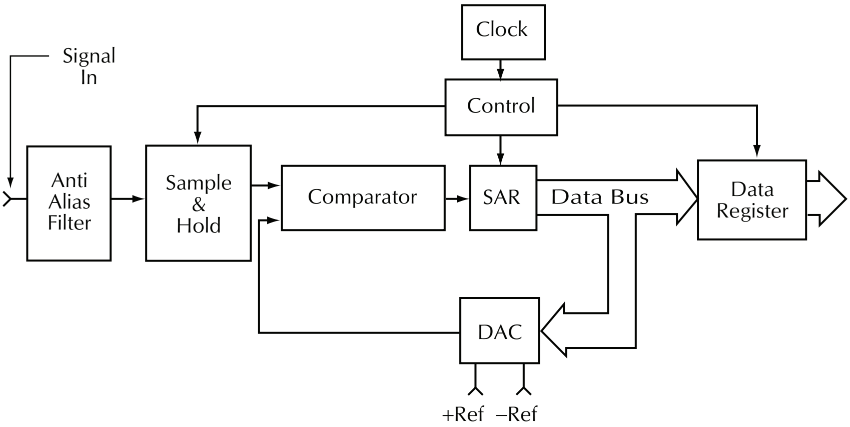


Figure 22-31. Successive approximation analog to digital converter.

have been tested. At the conclusion of all bit tests, the code in the DAC as well as the data register is within one LSB of the value equal to that representing the hold signal. The control circuit then issues an end of conversion signal that validates and maintains the status of the data register while clearing the DAC in preparation for the next conversion. The clocking rate for bit tests is necessarily greater than n times the audio signal sample rate.

In summary, an n-bit digital word is issued for each sample immediately following the end of the conversion period. The data in this form can be communicated directly to computer memory or to digital signal processing circuitry. The successive approximation technique does have at least two stringent requirements. An expensive high pole order analog antialiasing filter is required unless the sampling frequency is considerably greater than twice the maximum frequency to be passed in the signal spectrum. The resistors in the R–2R ladder network of the DAC must be very precise in order to insure linearity in the conversion process.

The delta-sigma analog to digital conversion technique is replacing the successive approximation technique in practically all audio applications. This has been made possible as a result of technological developments in the area of very large scale integrated circuits. Such circuits can now be made to successfully handle a mix of both digital and analog circuits. Delta-sigma analog to digital converters contain various combinations of elements depending upon the exact manner of employment. The basic element in all instances is a delta-sigma modulator. A simplified version of such a modulator is displayed in Fig. 22-32.

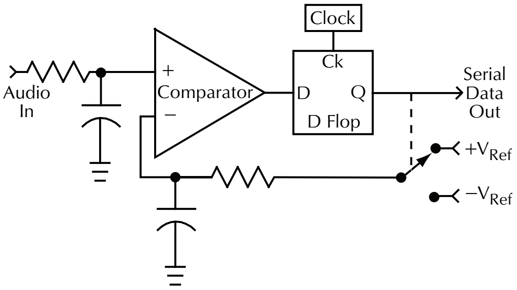


Figure 22-32. Simplified delta-sigma modulator.

The delta-sigma technique tracks the difference between the value of the audio signal at any given moment and a reconstructed audio signal that is derived through the summation of previous differences. If the audio signal at the moment of sampling exceeds the present value of the reconstructed signal, the output of the comparator is positive and the rising edge of the clock pulse generates a logic 1 at the output of the D flip flop. This in turn connects the RC integrator via an electronic switch to a positive voltage reference thus tending to increase the value of the reconstructed audio signal.

Had the reconstructed signal been greater than the audio signal at the moment of sampling, the comparator output would be negative forcing a logic zero at the output of the flip flop and a connection to a negative voltage reference. This would lead to a decrease in the value of the reconstructed audio signal. There will thus occur a serial bit stream at the output with 1s indicating level increases and 0s indicating level decreases. In order for the reconstructed signal to closely approximate the input audio signal the sampling rate must be much larger than that required to simply satisfy the Sampling Theorem. The delta-sigma technique thus employs oversampling with the oversampling rate often being 64×, 128×, or 192× the minimum required by the Sampling Theorem. The very high sampling rate greatly simplifies the design of the antialias filter with a simple RC lowpass often being sufficient. If the application is that of just signal delay, the serial data stream can be directly stored in memory and extracted after the required delay. The delayed serial data stream can be converted back to audio through the employment of a delta-sigma demodulator. A simplified circuit of such a demodulator is depicted in Fig. 22-33.

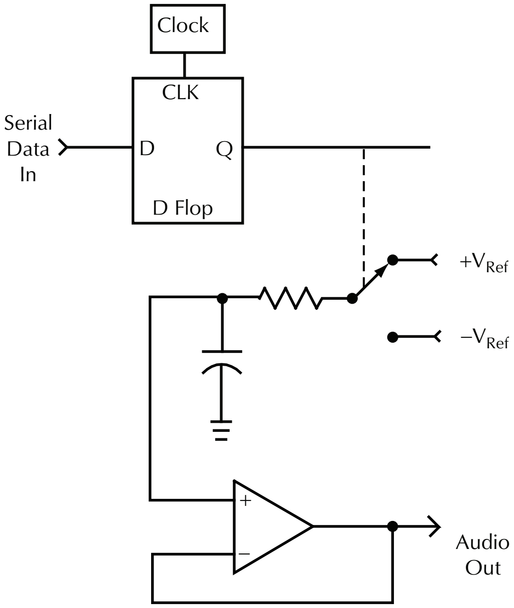


Figure 22-33. Simplified delta-sigma demodulator.

The operation of both the delta-sigma modulator and demodulator can be enhanced through the employment of multiple feedback loops of integration. In addition the voltage references can be dynamically adjusted through the employment of adaptive techniques with decisions being guided by the historical behavior of the serial bit chain. Alternatively, the application might require the conversion of the continuous serial bit stream into a succession of n-bit words at the normal sample rate.

In this instance, a special form of finite impulse response or FIR digital filter is required in addition to the modulator. This filter is termed a decimation filter. Decimation filters are usually constructed in stages so that a given filter might well consist of a cascade of several FIRs with downsampling by various factors of two occurring in each of the separate stages.

As an example suppose the objective is to produce a sequence of 16-bit samples at a sample rate of 48k samples per second. If the delta-sigma modulator oversamples at 64× then the required clock frequency would be 64 × 48k or 3.072 MHz. The serial bit stream emanating from the modulator at the clock rate is then fed into a suitably designed three stage FIR digital filter with decimation factors of 8, 4, and 2, respectively, the output from the filters being serial samples each consisting of 16 bits at a rate of 48k samples per second. Additionally the digital filter defines the passband of the system while offering a flat response from dc to beyond 20 kHz attended by an attenuation in the stop band of 90 dB or more.

Digital signal processing engines or DSPs generally require samples in parallel form. The final component then in the delta-sigma ADC might well be a 16-bit shift register for performing serial to parallel data conversion. In summary, the delta-sigma technique does not require any precision components as does the successive approximation technique. The high sampling rate makes possible the employment of a very simple analog antialias filter. In most instances there is no requirement for a sample and hold circuit and the decimation FIR filters precisely define the passband while offering large stopband rejection. The major shortcoming of the technique exists in the processing time required by the FIR filters. This processing time is usually of a few milliseconds.

At this juncture we will temporarily suspend our study of sampled data systems in order to return to a further pursuit of continuous time signals. The initial focus will be on system theory and a mathematical technique that treats both the transient and steady state response of physical systems under a single umbrella. This mathematical technique is the basis of description of analog filters and will serve as a guideline for developing a similar technique applicable to digital filters.

22.3 System Theory

The goal of system theory is to evolve a method of analysis such that certain classes of physical systems will conform to a signal flow of the form:

$$\text{Input or Excitation} \Rightarrow [\text{System}] \Rightarrow \text{Output or Response}$$

A slight rearrangement of which suggests the mathematical statement:

$$\text{Response function} = (\text{Excitation function})(\text{System function})$$

If this is truly an equality then it can be written:

$$\text{System function} = \frac{\text{Response function}}{\text{Excitation function}}$$

If the system function is to have a unique identity, related only to the system from which it stems, it cannot depend on either the response or the excitation functions just as a resistor which obeys Ohm's law does not depend on either the voltage or the current. The independence of the system function on the excitation function implies that neither the size of the excitation nor the time of its application is of importance. This requires that the system from which the system function is derived must be both linear and time invariant. In order to achieve the goal of system theory, it is necessary to transform the time domain equations governing the system into the corresponding equations in the complex frequency domain. Before exploring any of the details associated with these concepts mention should be made of the genius to whom we chiefly owe a great debt of gratitude for making it all possible.

Oliver Heaviside (1850-1925) was born in London and received only a secondary education. He was the nephew of Sir Charles Wheatstone but apparently was from the poor side of the family. He was employed as a youth by the Great Northern Telegraph Company but had to leave the company at age 24 because of encroaching deafness. Thereafter he set for himself the formidable task of studying and understanding Maxwell's treatise. Not only did he succeed in this endeavor he also proceeded to solve many of the outstanding electrical and communications problems of the time through the employment of analysis techniques which he had to evolve to compensate for his lack of formal mathematical training. He was a physicist in the highest sense of the word as he felt that physical ideas should be the master and mathematics the servant. In his view the servant must accede to the whims of the master. As a result, the highly placed academic mathematicians of his time railed against Heaviside's mathematical techniques and it was not until 1916 that they were finally proven to rest on firm mathematical foundations and were begrudgingly accepted by mathematicians.

Fortunately physicists and engineers had been using them in the meantime with great success. In 1887, barely 13 years after he began his serious studies, Heaviside solved the problem of distortionless transmission on telegraph and telephone cables and thus made long distance and wide bandwidth communication on these media a possibility. William Thomson had introduced serious study of this problem in 1855 but had not pursued it to its ultimate conclusion. Heaviside's solution to the cable problem is the first instance of electrical equalization. Heaviside's equalization technique was adopted by Professor Pupin at Columbia University in the United States and subsequently introduced to the American telephone industry. Heaviside published all of his work often at his own expense.

Heaviside, along with Professor Kennelly of Harvard University, was also the first to explain the ionosphere's role in long distance radio communication. Heaviside's contributions were well recognized and he received many honors toward the end of his life but his name is hardly ever referenced in modern textbooks. His mathematical techniques have been largely supplanted by the transform of Laplace, which is more general, or that of Fourier which is actually a special case of the two sided version of the Laplace transform. The systems approach employing principally physical reasoning was his alone and this uniquely warrants his continued memory.

In determining system or transfer functions the tool of choice is the Laplace transform. The Laplace transform is an integral transform that converts a function of the independent variable of time symbolized by t into some other function of the complex frequency variable symbolized by S where S has both real and imaginary parts

$$S = \sigma + j\omega \tag{22-97}$$

The two sided Laplace transform of a function of t is given by

$$F(S) = \int_{-\infty}^{\infty} f(t)e^{-St}dt \tag{22-98}$$

The Fourier transform is a special case of the Laplace transform wherein the frequency variable is purely imaginary.

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt \tag{22-99}$$

In analyzing systems, the transform tool to be employed is the one sided Laplace transform given by

$$F(S) = \int_{0}^{\infty} f(t)e^{-St}dt\,. \tag{22-100}$$

This is employed for two reasons. Firstly, the consideration of time signals defined only for $t \geq 0$ automatically satisfies the principle of causality. Secondly, the Laplace transform with complex frequency variable inherently takes into account any initial conditions which exist in the system at $t = 0$.

One can employ Eq. 22-100 in order to directly calculate the Laplace transform of a given $f(t)$. This is often not necessary, as tables are available containing the transforms of commonly encountered time functions. Such tables are called tables of function and transform pairs and can be found in most mathematical handbooks. Table 22-4 is an abbreviated table presented here for ready reference.

Table 22-4. A Short Table of Function Tranform Pairs

Function of Time	Transform
1	$\frac{1}{S}$
$\sin(\omega t)$	$\frac{\omega}{S^2+\omega^2}$
$\cos(\omega t)$	$\frac{S}{S^2+\omega^2}$
$\sinh(\gamma t)$	$\frac{\gamma}{S^2-\gamma^2}$
$\cosh(\gamma t)$	$\frac{S}{S^2-\gamma^2}$
t	$\frac{1}{S^2}$
$\frac{t^{n-1}}{(n-1)!}$	$\frac{1}{S^n}$
$e^{\gamma t}$	$\frac{1}{S-\gamma}$
$e^{-\gamma t}$	$\frac{1}{S+\gamma}$

Table 22-4. (cont.) A Short Table of Function Tranform Pairs

Function of Time	Transform
$e^{-\alpha t}\sin(\beta t)$	$\dfrac{\beta}{(S+\alpha)^2+\beta^2}$
$e^{-\alpha t}\cos(\beta t)$	$\dfrac{S+\alpha}{(S+\alpha)^2+\beta^2}$
$te^{-\alpha t}$	$\dfrac{1}{(S+\alpha)^2}$
$\dfrac{t^{n-1}e^{-\alpha t}}{(n-1)!}$	$\dfrac{1}{(S+\alpha)^n}$
$\dfrac{e^{-\alpha t}-e^{-\gamma t}}{\gamma-\alpha}$	$\dfrac{1}{(S+\alpha)(S+\gamma)}$
$\dfrac{1-\cos(\alpha t)}{\alpha^2}$	$\dfrac{1}{S(S^2+\alpha^2)}$

The laws of physics governing the behavior of physical systems mathematically appear usually in the form of time dependent differential or integro-differential equations. In the course of attaining the goal of system theory it is necessary to take the Laplace transform of each term in such equations. The mathematical operations involved are successive differentiation and or integration. Tables relating the time dependent mathematical operations and their respective transforms are also available. A brief such table will also be presented here for ready reference.

A few words of explanation of the entries in Table 22-5 are needed for clarification. In the first entry if $f(t)$ is the time domain function, its Laplace transform is represented by $F(S)$. In the second entry the time domain function is the first derivative of $f(t)$ with respect to t. The corresponding Laplace transform of this is S times the Laplace transform of $f(t)$ less the value of $f(t)$ at the instant of zero time. The third entry is the time domain function which represents the second derivative with respect to time of the function $f(t)$. The Laplace transform of this time dependent function is S squared times the transform of $f(t)$ less S times the initial value of $f(t)$and further less the value of the first time derivative of $f(t)$ evaluated at the initial instant. Finally, in the fourth entry, the time dependent function is the integral with respect to time of $f(t)$.

The transform of this time dependent function is the transform of $f(t)$ divided by S plus the time integral of $f(t)$ with respect to t evaluated at the initial instant. This last term probably requires even further explanation. For example, imagine that $f(t)$ represents an electrical current in a capacitor, then the integral of the current would be the electric charge on the capacitor and the value of the integral at the initial instant would be the charge on the capacitor when t is equal to zero.

Table 22-5. A Short Table of Operations Transform Pairs

Function of Real Variable	LaPlace Transform
$f(t)$	$F(S)$
$\dfrac{df(t)}{dt}$	$SF(S)-f(0)$
$\dfrac{d^2f(t)}{dt^2}$	$S^2F(S)-Sf(0)-\left[\dfrac{df(t)}{dt}\right]_{t=0}$
$\int f(t)\,dt$	$\dfrac{F(S)}{S}+\dfrac{1}{S}\int^{0} f(t)\,dt$

At this point it would be instructive to apply this method of analysis to a physical system commencing from first principles. So as not to be overly ambitious a relatively simple system will be chosen. This system will consist of a time dependent signal generator which has an open circuit emf denoted by $e_i(t)$ and an internal resistance of R_i. This generator is supplying a signal to the input circuit of a voltage amplifier whose input circuit can accurately be modeled as a resistance R_a in parallel with a capacitance C_a. The object is to determine the transfer function for this system where the excitation is $e_i(t)$, the exact form of which is arbitrary, and the response is to be $v_a(t)$ where $v_a(t)$ is the time dependent voltage appearing at the input terminals of the amplifier. The circuit thus appears as displayed in Fig. 22-34.

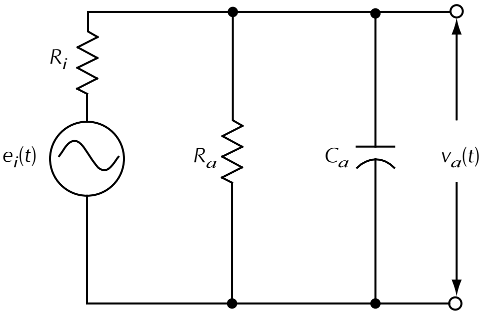


Figure 22-34. Amplifier input circuit model.

The next step is to apply the laws of physics by writing the differential equation which governs the system taking the time dependent voltage v_a as the dependent variable. The current in the resistor R_a is the voltage v_a divided by the resistance R_a. The current in the capacitor C_a is the time rate of change of the charge on this capacitor which in turn is the value of the capacitance multiplied by the time rate of change of the voltage across the terminals of the capacitor. The current in the resistor R_i is the sum of these currents. The emf of the generator is equal to the voltage across R_i plus the voltage v_a. The equation governing the circuit behavior is thus

$$e_i(t) = \left[\frac{v_a(t)}{R_a} + C_a\frac{dv_a(t)}{dt}\right]R_i + v_a(t) \qquad (22\text{-}101)$$

In this example, the initial charge on the capacitor at $t = 0$ will be taken to be zero although other choices are possible as determined by the prior history of the circuit before the generator is connected. The procedure continues by taking the Laplace transform of each term in this equation. As neither $e_i(t)$ nor certainly $v_a(t)$ are known explicitly as functions of time their Laplace transforms are implied by $E_i(S)$ and $V_a(S)$, respectively. The transformed equation becomes

$$E_i(S) = \left[\frac{V_a(S)}{R_a} + C_aSV_a(S)\right]R_i + V_a(S) \qquad (22\text{-}102)$$

Notice that the derivative with respect to time of a time dependent function has been replaced by multiplication of the transformed time function by the complex frequency variable S as indicated by the second entry in Table 22-5. There is also no contribution from initial conditions here because $v_a(t)$ is zero prior to $t = 0$. The transformed equation can be solved algebraically to obtain

$$V_a(S) = E_i(S)\left(\frac{\frac{1}{R_iC_a}}{S + \frac{R_a + R_i}{R_aR_iC_a}}\right) \qquad (22\text{-}103)$$

The result states that the response as a function of the complex frequency S is equal to the excitation expressed as a function of the complex frequency multiplied by some other function of the complex frequency S. This other function is the transfer function which will be denoted $H(S)$. It should be clear that

$$H(S) = \frac{\frac{1}{R_iC_a}}{S + \frac{R_a + R_i}{R_aR_iC_a}} \qquad (22\text{-}104)$$

The transfer function is really a powerful code that allows one to determine how the system will behave regardless of the nature of the excitation. In fact, as will now be shown, the transfer function is the description in the complex frequency plane of the system's response to an impulse applied as an excitation in the time domain and the inverse Laplace transform of the transfer function is the time domain description of such response.

What is a unit impulse and what is its Laplace transform? The concept of impulse originated in classical mechanics wherein Newton's second law of motion states the force applied to a particle equates to the time rate of change of the particle's linear momentum or

$$F = \frac{d}{dt}(mv) \qquad (22\text{-}105)$$

In Eq. 22-105, m is the particle mass, v is the particle's instantaneous velocity, and the product of the two is the particle linear momentum. Now, given a force which itself may be time dependent what change in the particle's momentum occurs in some definite time, say t_0? The answer is

$$\int_0^{t_0} F(t)dt \qquad (22\text{-}106)$$

This integral is called the impulse of the force or more simply just the impulse. The result of the impulse is a definite change in momentum of the particle. However a given change in momentum can be brought about in a variety of ways. The force may be weak but last for a long period of time or it may be strong but last for just a short time. As long as the integrals are equal in the two cases the result will be the same. Now consider a time function which has a time dependence as displayed in Fig. 22-35.

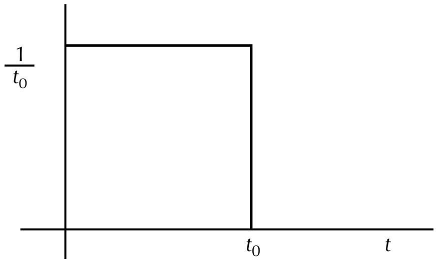


Figure 22-35. Time dependent function.

First of all note that the area under the curve is unity. Now imagine that t_0 is allowed to become increasingly smaller while the height of the plot which is inversely proportional to t_0 becomes increasingly larger with the total area remaining at unity. In the limit as t_0 approaches 0 the height of the plot approaches ∞ while the area remains at unity. The impulse of this function, that is the area under the curve, is always unity and hence it is called a unit impulse but the function being plotted versus time has grown infinitely large while its existence has become infinitesimally short. This peculiar function is called the impulse function or $\delta(t)$. The Laplace transform of this function can be found through the steps of the limiting process used in arriving at the function itself as follows.

$$\Delta(S) = lim_{t_0 \to 0}\int_0^{t_0}\frac{1}{t_0}e^{-St}dt \tag{22-107}$$
$$= lim_{t_0 \to 0}\frac{1-e^{-St_0}}{St_0}$$

In evaluating this limit the term e^{-St_0} is replaced by the infinite series expansion for this quantity and only the leading terms of the series are retained because all of the higher order terms are negligibly small compared with the leading terms. Therefore,

$$\Delta(S) = lim_{t_0 \to 0}\frac{1-1+St_0}{St_0} \tag{22-108}$$
$$= 1$$

The conclusion is that the Laplace transform of $\delta(t)$ is 1! Note that the dimension of the ordinate in Fig. 22-35 is reciprocal time and that the area is dimensionless. Now if a physical quantity such as voltage has an impulsive behavior this is represented by $k\delta(t)$ where the size of k represents the value of the voltage and the dimensions of k are volt seconds. The product $k\delta(t)$ is termed an impulse of strength k. Recall that for the simple system under study in Eq. 22-109

$$V_a(S) = E_i(S)\left(\frac{\frac{1}{R_iC_a}}{S+\frac{R_a+R_i}{R_aR_iC_a}}\right) \tag{22-109}$$

If the generator exciting the system has a time behavior of an impulse function of strength k, this equation will become

$$V_a(S) = k\left(\frac{\frac{1}{R_iC_a}}{S+\frac{R_a+R_i}{R_aR_iC_a}}\right) \tag{22-110}$$

Alternatively, Eq. 22-109 may be written

$$V_a(S) = kH(S) \tag{22-111}$$

It is possible to explore also what the system's behavior is in the time domain for such an excitation. This requires taking the inverse Laplace transform of the frequency domain behavior that is

$$[kH(S)]^{-1} = kh(t) \tag{22-112}$$

In Eq. 22-112, the operation of taking the inverse Laplace transform is indicated on the left with the corresponding time domain function on the right. The mathematical operations of taking the Laplace transform and its inverse are linear and hence

$$[H(S)]^{-1} = h(t) \tag{22-113}$$

The function $h(t)$ is the system's time domain response to a unit impulse. Eq. 22-113 tells one that $h(t)$ and $H(S)$ constitute a function transform pair. This implies that the system transfer function is the system's response expressed in the complex frequency domain to a unit impulse applied in the time domain. There is a formal way for taking the inverse Laplace transform but it is difficult as it involves a contour integration in the complex plane. The procedure which is usually employed is to consult a table of function, transform pairs, and hopefully find an appropriate $f(t) \Leftrightarrow F(S)$ combination. Pair number nine in Table 22-4 is appropriate in this instance when one identifies

$$\gamma = \frac{R_a+R_i}{R_aR_iC_a} \tag{22-114}$$

and recognizes that $1/R_iC_a$ is just a constant.

The impulse response of this simple system as a function of time is then

$$kh(t) = \frac{k}{R_iC_a}e^{-\frac{(R_a+R_i)t}{R_aR_iC_a}} \tag{22-115}$$

Graphically this has the form of a decaying exponential.

In words, what has happened is this. The generator at $t = 0$ produces an impulse of strength k implying an infinitely large voltage for an infinitesimally short period of time. Subsequently, the voltage at the amplifier terminals goes from 0 to a *finite value* in an infinitesimally short period of time. This voltage then decays exponentially with time from this finite value and approaches zero asymptotically as t grows infinitely large. This is the time domain description but in order to learn all of this it was necessary to transform the problem to the complex frequency domain. What then does it look like in this domain? For simplicity let R_i and R_a each equal 1 Ω and let $C_a = 1$ F. An unlikely set of values but nevertheless convenient. With this assignment,

$$H(S) = \frac{1}{S+2} \tag{22-116}$$

Recall that in general $S = \sigma + j\omega$ that is, $H(S)$ is complex.

$$H(S) = \frac{1}{\sigma + j\omega + 2} \tag{22-117}$$

This can be written in the form

$$H(S) = |H|e^{j\varphi} \tag{22-118}$$

where,
$|H(S)|$ is the absolute magnitude of $H(S)$,
φ is the phase of $H(S)$.

$$|H| = \frac{1}{\sqrt{(\sigma+2)^2 + \omega^2}} \tag{22-119}$$

$$\varphi = \operatorname{atan}\left(\frac{-\omega}{\sigma+2}\right) \tag{22-120}$$

$|H|$ is a two dimensional surface which has the appearance in Fig. 22-36.

In Fig. 22-36 the plane below the surface is the complex frequency or S plane. The transfer function surface exists in three dimensions so that each point on the surface has in general x, y, and z coordinates. The x axis corresponds to σ, the y axis corresponds to $j\omega$, and the z axis represents the magnitude of $H(S)$. At the point $\sigma = -2$, $\omega = 0$, the surface rises to infinity as the magnitude of the transfer function is infinite at that point. This singular point is called a pole of the transfer function. The location of the pole in the left half of the complex plane, brought about by σ being negative, is very significant as this is what led to the impulse response being a decaying exponential. If the pole were located in the right half of the complex plane, the impulse response would be a growing exponential and the system would not be a physically realizable stable system. This surface is sufficient to describe the complete behavior of this system for all types of excitations in both the steady state as well as the transient state. For example, if one takes a slice through this surface by assigning σ the value of zero what appears is the steady state amplitude response of the system for sinusoidal excitation. This is depicted in Fig. 22-37.

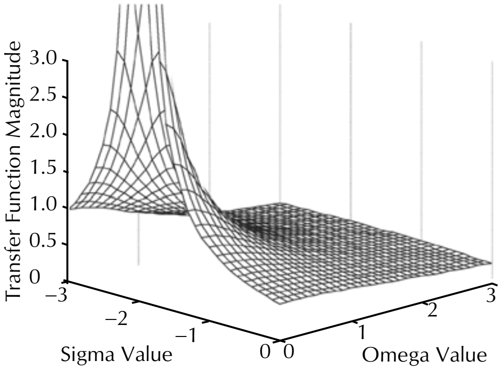


Figure 22-36. Transfer function surface.

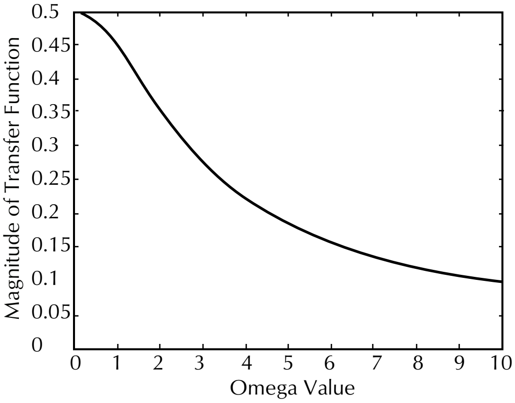


Figure 22-37. Amplitude response for sinusoidal excitation.

Finally a slice taken through the point $\sigma = -2$ and an arbitrary value of ω makes an angle with the real axis which is $\varphi = \operatorname{atan}(-\omega/2)$. This is the phase response of the system for steady state sinusoidal excitation.

One other possible feature of transfer functions requires examination. In order to see this, the original circuit needs to be modified by the insertion of a capacitor between the generator and the amplifier of the original example. Let this capacitor also for simplicity have a value of 1 F. The new transfer function will become

$$H(S) = \frac{S}{S^2 + 3S + 1} \tag{22-121}$$

and the new transfer function surface is depicted in Fig. 22-38.

The transfer function surface now exhibits two poles rather than the single one of the former case. This is because the denominator of the transfer function is now a quadratic having two distinct roots or values of S for which the denominator becomes zero and the magnitude of the transfer function becomes infinite. The really new feature, however, is

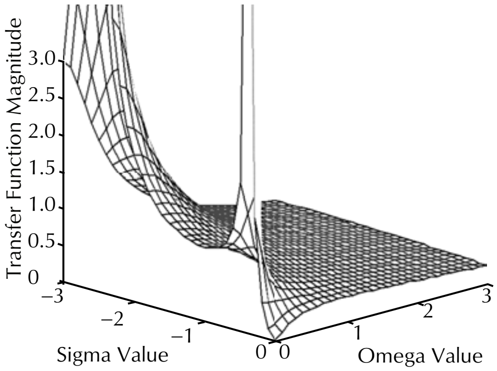


Figure 22-38. Modified transfer function surface.

the point at which the surface contacts the S plane. This point denotes a value for S at which the transfer function has a value of zero. In this instance, this occurs when both σ and ω are equal to zero. This point is called a zero of the transfer function. The shape of a general transfer function surface is determined by the location of the poles and zeros. The fact that the zero in this instance occurs at the origin in the S plane means that the system has no response to direct current in the steady state. Dc of course means that the exciting frequency is zero. The steady state amplitude response with sinusoidal excitation for this modified system is displayed in Fig. 22-39.

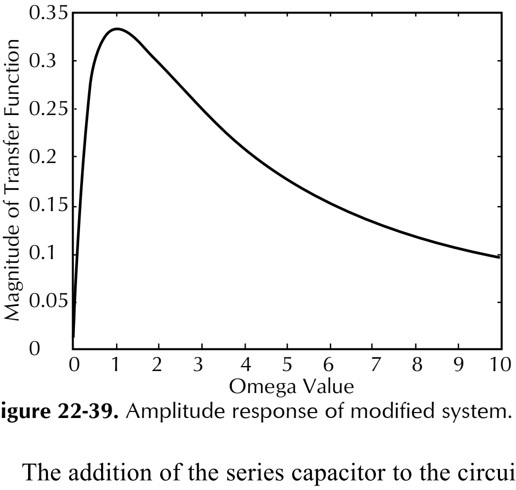


Figure 22-39. Amplitude response of modified system.

The addition of the series capacitor to the circuit has produced a system which now constitutes a bandpass filter which should be evident from the shape of the amplitude response curve. The close linkage between the shape of the transfer function surface and the locations of the poles and zeros of the transfer function have led to a rapid analysis of transfer functions termed pole-zero analysis. This technique should be explored in detail as it is not computationally intensive and it yields most of the important answers with regards to system behavior.

22.3.1 Pole and Zero Analysis

The circuit diagram for the modified amplifier input circuit is displayed in Fig. 22-40.

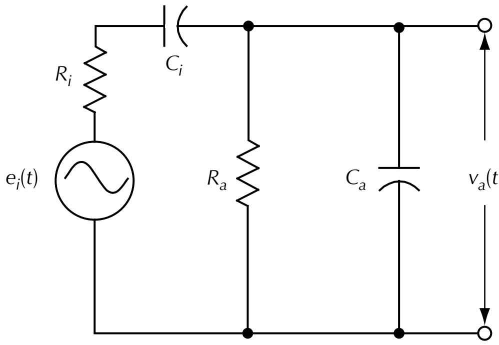


Figure 22-40. Modified amplifier input circuit.

If one were analyzing this circuit by employing the techniques presented in Chapter 8 *Interfacing Electrical and Acoustic Systems*, the ratio of the output voltage to the input voltage would be written as

$$\frac{V_a}{E_i} = \frac{\dfrac{R_a\dfrac{1}{j\omega C_a}}{R_a + \dfrac{1}{j\omega C_a}}}{R_i + \dfrac{1}{j\omega C_i} + \dfrac{R_a\dfrac{1}{j\omega C_a}}{R_a + \dfrac{1}{j\omega C_a}}} \tag{22-122}$$

If the capacitors are uncharged and there are no currents prior to the connection of the generator at $t = 0$, Eq. 22-121 can be put easily into the language of the Laplace transform simply by replacing $j\omega$, wherever it appears, by S. In doing so, it is important to remember that S still represents $\sigma + j\omega$. After this substitution and some algebraic simplification, the recasted equation becomes

$$\frac{V_a(S)}{E_i(S)} = \tag{22-123}$$

$$\frac{\dfrac{1}{R_iC_a}S}{S^2 + \dfrac{(R_aC_a + R_iC_i + R_aC_i)}{R_aC_aR_iC_i}S + \dfrac{1}{R_aC_aR_iC_i}}$$

At this point one may substitute typical values for each of the circuit components and deal with the collection of numbers that results or pick some values that will illustrate the procedure to be followed without requiring tedious numerical calculation. We will pursue the latter course by taking unit values for each of the components in which case Eq. 22-122 becomes simply

$$\frac{V_a(S)}{E_i(S)} = \frac{S}{S^2+3S+1} \tag{22-124}$$

The objective at this point is to locate the poles and zeros of the transfer function in the complex *S* plane. By inspection, if *S* is set equal to zero, the value of the transfer function also becomes zero. Therefore, there is a single zero located at the origin. The poles of the transfer function correspond to values of *S* which when substituted in the denominator force the denominator to have a value of zero and thus forcing the transfer function to become infinite. The denominator here is a quadratic in *S* and hence there are two values of *S* which make the denominator zero therefore this transfer function has two poles. The pole locations are identified by factoring the denominator. A straightforward application of the quadratic formula leads to

$$\frac{V_a(S)}{E_i(S)} = \frac{S}{(S+0.382)(S+2.618)} \tag{22-125}$$

The poles, then, are located at $S = -0.382$ and $S = -2.618$. The pole and zero diagram depicts these locations and is presented in Fig. 22-41.

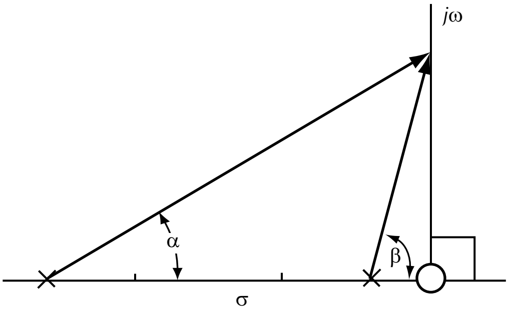


Figure 22-41. Pole-zero diagram.

The location of the zero in the diagram of Fig. 22-8 is indicated by the circle at the origin. A line drawn from this location to an arbitrary value of ω is indicated in the figure. This line makes a right angle with the positive real or σ axis. The locations of the two poles are indicated by *x*s. In this instance, both of the poles have real values and each of them is located on the negative real axis. Lines are drawn from each of the poles to the selected arbitrary value of ω and these lines make the angles α and β, respectively, with the positive real axis. The phase response function can be readily calculated by observing the angles in this diagram. In the case of steady state sinusoidal excitation, the phase response function is simply the phase difference between the output signal and the input signal expressed as a function of ω. This phase difference can be extracted directly from the figure. This phase difference is the sum of all angles contributed by zeros less the sum of all angles contributed by poles. In this instance there is only a single zero and its angle is π/2. There are two poles providing collectively the angles α and β. Upon denoting the phase difference by φ, the phase response function becomes

$$\begin{aligned}\varphi(\omega) &= \frac{\pi}{2} - \alpha(\omega) - \beta(\omega) \\ &= \frac{\pi}{2} - \operatorname{atan}\left(\frac{\omega}{2.618}\right) - \operatorname{atan}\left(\frac{\omega}{0.382}\right)\end{aligned} \tag{22-126}$$

The amplitude response for sinusoidal excitation is the magnitude of the output signal divided by the magnitude of the input signal expressed as a function of ω. This may be calculated directly from either Eqs. 22-123 or 22-124. One simply substitutes *j*ω for *S* wherever *S* appears and then determines the magnitude of the resulting complex quantity. This substitution yields

$$\frac{j\omega}{-\omega^2+3j\omega+1} \tag{22-127}$$

The magnitude of the numerator is just ω. Recall that the magnitude of the denominator is the square root of the sum of the squares of the real and imaginary part hence, the amplitude response function is written as

$$\left|\frac{V_a}{E_i}\right| = \frac{\omega}{\sqrt{\omega^4+7\omega^2+1}} \tag{22-128}$$

Judging from the examples presented so far, one might have the false impression that transfer function poles are always real quantities. In fact, the poles are more often complex. As an illustration of complex poles, an example will be drawn from a real problem dealing with a magnetic phonograph cartridge. The passive circuit elements associated with the cartridge's construction and the input circuit of the preamplifier to which it may be connected appear in Fig. 22-42.

The internal structure of the cartridge is modeled as an ideal voltage source in series with the

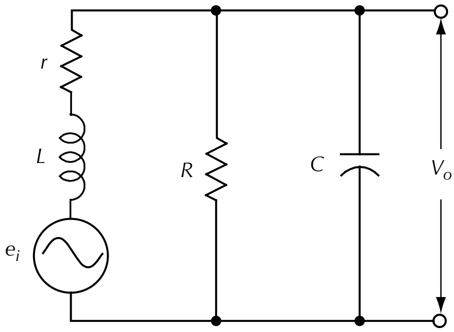


Figure 22-42. Magnetic phonograph cartridge circuit.

inductance of the cartridge's coil represented by *L* and the winding resistance of the coil represented by *r*. The resistive load presented by the preamplifier to the cartridge is *R* and the shunt capacitance of the preamplifier and the associated connecting cable is represented by a total capacitance *C*. The exact expression for the transfer function is

$$\frac{V_0(S)}{E_i(S)} = \frac{\frac{1}{LC}}{S^2 + \frac{\left(\frac{L}{R} + rC\right)}{LC}S + \left(\frac{r+R}{R}\right)\frac{1}{LC}} \qquad (22\text{-}129)$$

In practice, the load resistance *R* is set to be about fifty times as large as the cartridge resistance, *r*. Furthermore, *L*/*R* >> *rC* for typical systems, so little error is made in using the following simpler form

$$\frac{V_0(S)}{E_i(S)} \approx \frac{\frac{1}{LC}}{S^2 + \left(\frac{1}{RC}\right)S + \frac{1}{LC}} \qquad (22\text{-}130)$$

The analysis here is not concerned with recording characteristics, pre-emphasis, or de-emphasis but rather with the faithful transmission of the signal produced by the generator regardless of its form. There are no zeros in the transfer function of Eq. 22-129. This implies that there is response extending to dc or zero frequency. There are two poles as evidenced by the denominator being a quadratic in the complex frequency variable. These facts taken together define the response to be that of a second order low pass filter where the shape of the response depends solely on the locations of the poles. The pole order is dictated by the highest power of the complex frequency variable that appears in the polynomial that constitutes the denominator of the transfer function. Regardless of the placement of the poles in the left half of the complex plane, each pole contributes an asymptotic attenuation rate of 6 dB/octave and an asymptotic phase shift of $-\pi/2$ radians. The attenuation rate of this filter at high frequencies is thus 12 dB/octave with an attendant phase shift at high frequencies of π radians. The question becomes where should the poles be placed in order to obtain ideal response and what constitutes "ideal" response. In studying this placement, it is useful to cast the transfer function in a standard form which is

$$\frac{V_0(S)}{E_i(S)} = \frac{{\omega_0}^2}{S^2 + \alpha\omega_0 S + {\omega_0}^2} \qquad (22\text{-}131)$$

The behavior of the polynomial in the denominator of Eq. 22-131 now hinges on the significance assigned to the value of ω_0 and the value of the coefficient α. There are a multitude of possibilities from which three distinct classes of polynomials have been found to possess useful characteristics. These are the Bessel, Butterworth, and Chebyshev polynomials.

The Bessel polynomials offer the most linear phase behavior with the reciprocal of ω_0 being equal to the signal group delay at zero frequency.

The Butterworth polynomials offer maximally flat amplitude response with ω_0 being the angular frequency at cut-off, i.e., the angular frequency at which the amplitude response is −3 dB for a low or high pass filter. Note, however, it corresponds to the center angular frequency for a bandpass filter.

The Chebyshev polynomials feature controlled ripples in the amplitude response while affording a rapid attenuation rate in the vicinity of the filter cut-off point. The value of ω_0 for the Chebyshev corresponds to the angular frequency at which the attenuation at the cut-off edge of the pass band is equal to the ripple minimum in the pass band.

If the "ideal" to be achieved is that of flattest amplitude response, the polynomial of choice will be a Butterworth. An interesting property of the Butterworth polynomials is that regardless of the polynomial order, the poles are located on a semi-circle of radius ω_0 in the left half of the *S* plane. Furthermore, for the second order Butterworth polynomial, which is the case here, α has a value of $\sqrt{2}$. These pole locations are depicted in Fig. 22-43.

In Fig. 22-43 the radius of the semi-circle is ω_0, the poles are complex conjugates, and they are separated by an angular interval of $\pi/2$. If the denominator had been third order, there would have been three poles. One would be on the negative real axis at $-\omega_0$. The other two would form a conjugate pair placed such that the angle between a given pole and its nearest neighbor would be $\pi/3$. The pattern

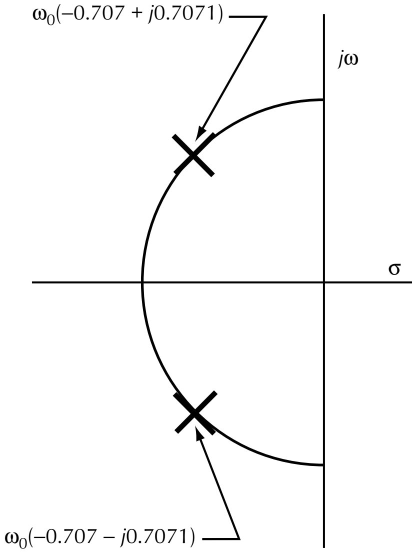


Figure 22-43. Complex poles of second order Butterworth polynomial.

continues in this fashion for all higher orders with odd orders giving a single pole on the negative real axis with attendant conjugate pairs. The even orders contribute only conjugate pairs and for all order n, the angle between any pole and its neighbor is π/n.

The analysis of the phonograph cartridge can now be completed by comparing Eqs. 22-130 and 22-131 while requiring α to be $\sqrt{2}$. This comparison leads to the identifications

$$\begin{aligned}\omega_0^2 &= \frac{1}{LC}\\ \frac{1}{RC} &= \sqrt{2}\omega_0\end{aligned} \qquad (22\text{-}132)$$

The next step is to solve Eq. 22-131 for C to obtain

$$C = \frac{L}{2R^2} \qquad (22\text{-}133)$$

A typical high quality cartridge has an inductance internal to its structure of 0.5 H. Additionally the load resistance R is usually set at $47\,\mathrm{k}\Omega$. When these values are substituted into Eq. 22-133, it is found that the total allowable shunt capacitance in order to obtain a maximally flat response is

$$\begin{aligned}C &= \frac{0.5}{2(47{,}000)^2}\\ &= 113\text{ pF}\end{aligned}$$

To find the cutoff frequency, use

$$\begin{aligned}f_0 &= \frac{\omega_0}{2\pi}\\ &= \frac{1}{2\pi\sqrt{LC}}\end{aligned} \qquad (22\text{-}134)$$

Using Eq. 22-134, and the value of shunt capacitance in the example, the cut-off frequency for the cartridge circuit becomes

$$\begin{aligned}f_0 &= \frac{1}{2\pi\sqrt{LC}}\\ &= 21\text{ kHz}\end{aligned}$$

This value of the cut-off frequency is acceptable but the low value of C forces careful treatment of the preamplifier circuit and the phonograph cartridge connecting cable. Fig. 22-44 displays the amplitude response employing the values found in the above example. The plot is in the form of $20\,\mathrm{dB}\log|V_0(S)/E_i(S)|$ vs. $\log f$.

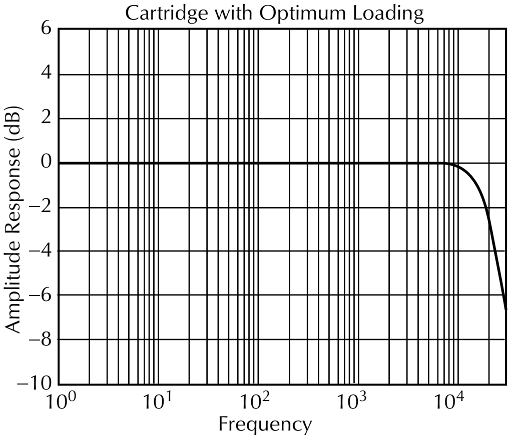


Figure 22-44. Amplitude response of optimally loaded phonograph cartridge.

In the event the shunt capacitance is larger than that required for the Butterworth, the response will exhibit a peak as illustrated in Fig. 22-45.

Finally, Fig. 22-46 shows the result of working with the exact form of the transfer function as expressed in Eq. 22-129. In this instance, the L and r values are those supplied by the manufacturer of a well-esteemed cartridge, the R value is taken as $47\,\mathrm{k}\Omega$, and C is calculated as before for optimum loading.

A comparison of Figs. 22-44 and 22-46 indicates that the error made in neglecting the internal resistance of the cartridge is only a fraction of a decibel and hence is insignificant.

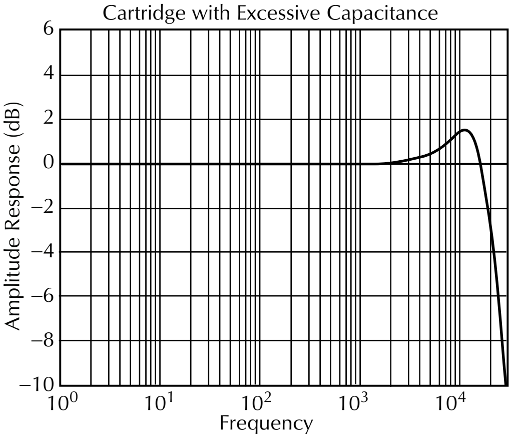


Figure 22-45. Amplitude response with excessive shunt capacitance.

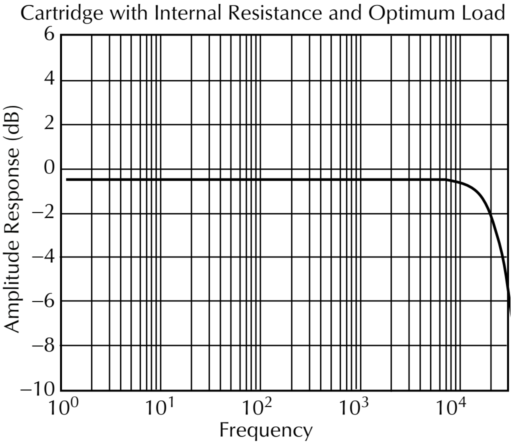


Figure 22-46. Response employing exact form of transfer function.

In order to completely describe a system's steady state frequency response with sinusoidal excitation one must describe the amplitude response as a function of frequency or angular frequency and, additionally, must describe the phase response as a function of frequency or angular frequency. This information is all contained in the complex transfer function and can be separated out by examining separately the magnitude of the transfer function versus the chosen frequency variable and the angle of the transfer function versus the chosen frequency variable. Bode plots are a standard way of displaying these results. The amplitude response as displayed in a Bode plot is in the form of 20 dB times the log to the base ten of the magnitude of the transfer function versus log to the base ten of the chosen frequency variable. The phase response as displayed in a Bode plot is in the form of the phase difference between the output and input, which is the phase of the transfer function, plotted versus the log to the base ten of the chosen frequency variable. Figs. 22-47 and 22-48 are the Bode plots for the phonograph cartridge with optimum shunt capacitance.

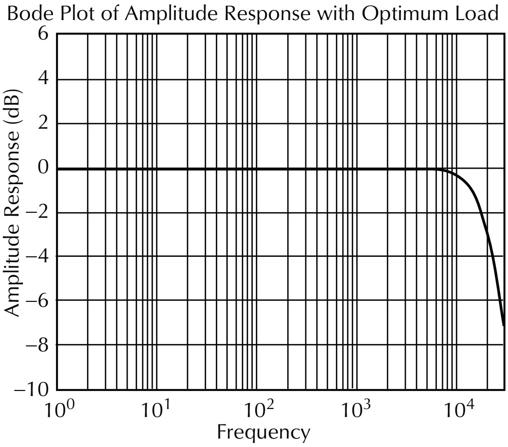


Figure 22-47. Bode plot of cartridge amplitude response.

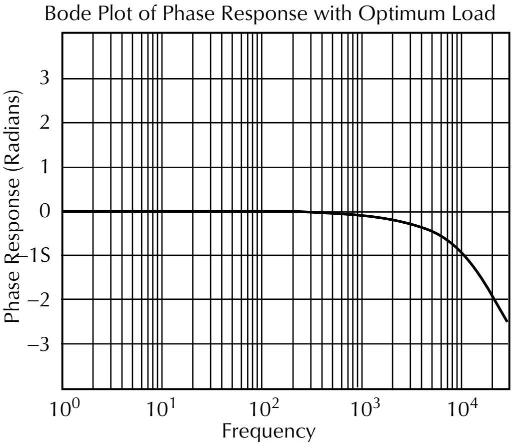


Figure 22-48. Bode plot of cartridge phase response.

22.3.2 Further Considerations

From the work thus far one might draw the mistaken conclusion that transfer functions are always dimensionless. This has been the case in the cited examples as both the excitation and the response have been voltages. This is certainly not always the case. In electrical systems the excitation may be a voltage or a current while the response may be a current or a voltage. For those instances where the excitation and response are similar quantities, the transfer function will be dimensionless. In the other instances the dimensions will be those of an admittance or an impedance. More particularly, however, would be those instances where the system under study is a transducer such as a microphone or a loudspeaker. The excitation for a microphone is an acoustic pressure and the response is a voltage. For a loudspeaker just the converse is true, the excitation is a voltage and the response is an acoustic pressure. Thus the transfer function would have the dimensions of volt per pascal (V/Pa) or pascal per volt (Pa/V), respectively. In any event, the transfer function can be written in the form of a scaling factor, which includes numerical sensitivity as well as the appropriate dimensions, multiplied by a frequency dependent term that is itself dimensionless.

$$H(S) = (\text{Scale factor with dimensions}) \\ (\text{Dimensionless frequency dependent function}) \tag{22-135}$$

In the foregoing, much has been said with regard to allowable pole locations whether real or complex. A physically realizable stable system is one whose impulse response decays with time. In order for this to be true each pole in the transfer function must have a negative real part. Geometrically, this means that all poles must be located in the left half of the S plane. Given that this is the case, are there similar restrictions with regard to the locations of the zeros of transfer functions? The answer to this question is in the negative. Zeros whether real, imaginary, or complex can be located anywhere without any influence on stable system realizability. However, zero locations do play a significant role with regard to a system's phase response.

System phase responses can be divided into the two categories of minimum phase systems and non-minimum phase systems. Physically stated, a minimum phase system is one that can release its stored energy in a minimum time. When viewed in the frequency domain, however, minimum time translates to minimum phase shift. This may be explored by studying the impulse transient response in the time domain or, alternatively, by comparing the phase response in the frequency domain. Consider two non-identical systems which possess the same amplitude response but differ in their phase responses. Can one construct such systems and, if so, how do their pole-zero diagrams differ? To show that such systems can be constructed, examine the following two transfer functions.

$$H_m(S) = k\frac{S+3}{(S+1)(S+2)} \tag{22-136}$$

$$H_n(S) = k\frac{S-3}{(S+1)(S+2)} \tag{22-137}$$

In Eqs. 22-136 and 22-137, k is the same constant. The amplitude response from Eq. 22-136 is

$$|H_m(S)| = k\frac{\sqrt{\omega^2+3^2}}{\sqrt{\omega^2+1^2}\sqrt{\omega^2+2^2}}$$

The amplitude response of Eq. 22-137 is

$$|H_n(S)| = k\frac{\sqrt{\omega^2+(-3)^2}}{\sqrt{\omega^2+1^2}\sqrt{\omega^2+2^2}}$$

Even just a cursory examination of the amplitude response expressions reveals their equality. The difference between Eq. 22-136 and Eq. 22-137, then, must be revealed in the phase responses. The pole-zero diagrams for these transfer functions are depicted in Fig. 22-49.

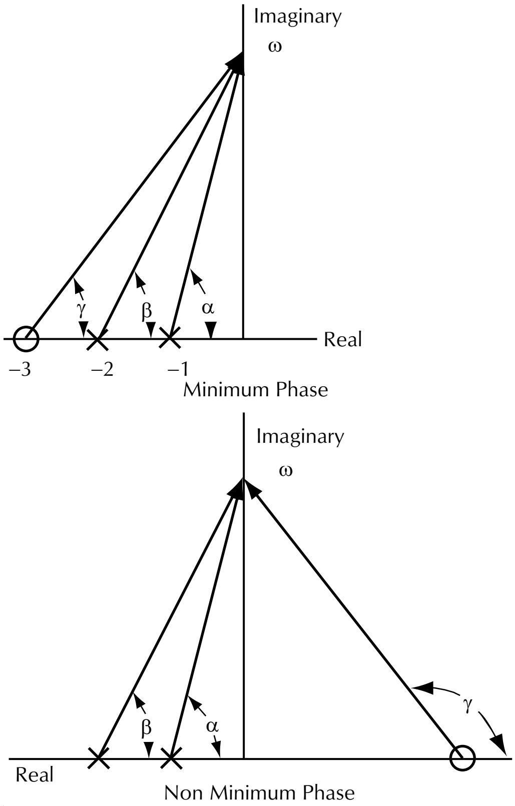


Figure 22-49. Pole-zero diagram comparison.

Eq. 22-136 describes a minimum phase system and from Fig. 22-49 its phase response is

$$\begin{aligned}\varphi_m &= \gamma - \alpha - \beta \\ &= \operatorname{atan}\left(\frac{\omega}{3}\right) - \operatorname{atan}\left(\frac{\omega}{1}\right) - \operatorname{atan}\left(\frac{\omega}{2}\right)\end{aligned} \tag{22-138}$$

Eq. 22-137 describes a non-minimum phase system and, again from Fig. 22-49, its phase response is

$$\begin{aligned}\varphi_n &= \gamma - \alpha - \beta \\ &= \pi - \operatorname{atan}\left(\frac{\omega}{3}\right) - \operatorname{atan}\left(\frac{\omega}{1}\right) - \operatorname{atan}\left(\frac{\omega}{2}\right)\end{aligned} \tag{22-139}$$

In writing Eq. 22-139 it must be noted that angles are always drawn relative to the sense of the positive real axis and hence the angle γ is 180° or π less the angle included at the base of the triangle on the right which is $\operatorname{atan}(\omega/3)$. These two phase responses are displayed in Fig. 22-50.

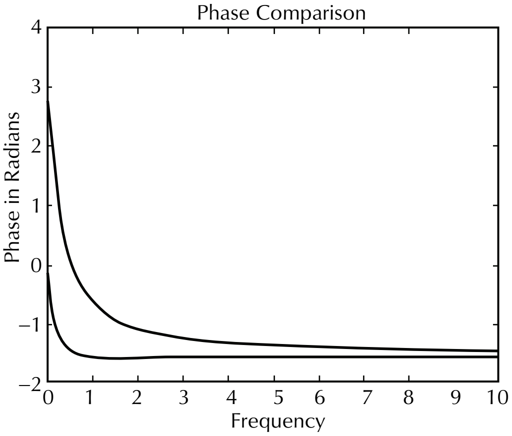


Figure 22-50. Non-minimum and minimum phase comparison.

In Fig. 22-50 φ_n is greater than φ_m at all finite frequencies and only approaches equality with φ_m as the frequency becomes infinite. The conclusion to be drawn is that minimum phase systems are characterized by having their transfer function zeros located in the left half of the S plane or, in worst case, on the imaginary axis. The importance of a system being of the minimum phase type lies in the fact that only minimum phase systems can be equalized. The validity of this statement, although not immediately obvious, will become apparent in the next section.

22.3.3 Equalization

Equalization as applied to sound reinforcement systems refers to the process of electrically modifying the signal being fed to the system so as to correct for some anomaly or anomalies exhibited by the system's natural behavior. Most systems are built up of several subsystems, each of which have individual transfer functions when considered as stand alone devices. Fig. 22-51 is suggestive of such a case.

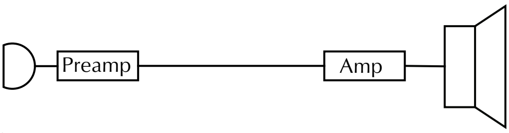


Figure 22-51. Simple reinforcement system.

The overall transfer function of the system will be the product of the individual transfer functions considered as stand alone devices under two circumstances. Each device is insensitive to source impedance and load impedance or the individual transfer functions have been determined subject to the appropriate source and load impedances. Assuming this to be the case,

$$H_{overall} = H_{mic}H_{preamp}H_{amp}H_{loudspeaker}$$

This will be of the general form

$$H_{overall} = k\frac{(S+a)(S+b)(S+c)\ldots}{(S+\alpha)(S+\beta)(S+\gamma)(S+\delta)\ldots} \tag{22-140}$$

In Eq. 22-140, a, b, c, etc. and α, β, γ, δ, etc. are constants which define the locations of the zeros and poles, respectively. If $H_{overall}$ displays anomalies in regards to either its amplitude or phase characteristic, a procedure exists for improving the overall results under a restricted set of circumstances. This procedure amounts to canceling a troublesome zero by a superimposed pole and canceling a troublesome pole by a superimposed zero. This procedure is called pole-zero compensation or otherwise known as equalization. The picture now becomes as in Fig. 22-52.

The equalizer's transfer function must have zeros for the troublesome poles and poles for the troublesome zeros. Recall that for physically realizable systems, however, the poles must be in the left half of the S plane. Therefore if the original system happened to have a troublesome zero in the right half of the S plane, which it could have if the anomaly

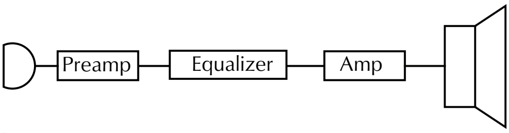


Figure 22-52. Simple reinforcement system with equalizer.

were non-minimum phase, it would be *impossible* to cancel this zero by a superimposed pole. Therefore such a system could not be truly equalized. If one were to attempt to experimentally equalize a non-minimum phase system by, say, flattening its amplitude response, this process would introduce violent behavior in the phase response and vice versa. The process of equalization is restricted to minimum phase shift systems. When equalization is applied to correct anomalies in minimum phase systems, the correction in amplitude and phase behavior go hand in hand and will occur simultaneously. Equalization should only be applied to minimum phase systems and even when it can be employed, it must be employed with care so as not to force any subsystem beyond its range of linear operation.

As an instance of a properly applied equalization consider the following example. Fig. 22-53 displays the amplitude response curve of a wide band loudspeaker which has a single resonance centered on 1 kHz.

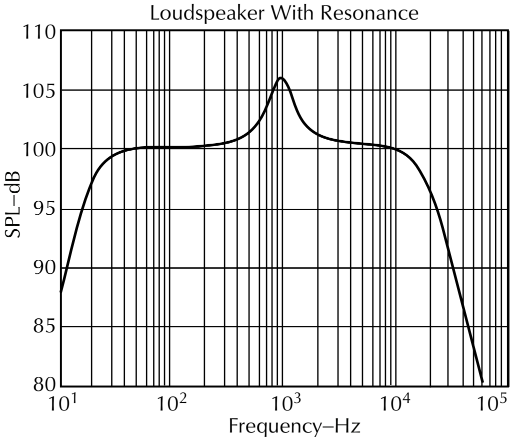


Figure 22-53. Loudspeaker with resonance anomaly.

The transfer function describing this system can be written as the product of a transfer function describing a loudspeaker without a resonance, H_l, with a transfer function which describes the anomalous resonance behavior, H_r, as given in Eq. 22-141.

$$H = H_l H_r \tag{22-141}$$

In this particular case, the resonance produces a 6 dB peak in the amplitude response at the resonant frequency. A transfer function which matches the peak and overall shape of the resonance is given by

$$H_r = \frac{S^2 + \omega_0 S + \omega_0^{\ 2}}{S^2 + 0.5\omega_0 S + \omega_0^{\ 2}} \tag{22-142}$$

Writing Eq. 22-142 in the factored form as presented in Eq. 22-140 more readily identifies the pole and zero locations.

$$H_r = \frac{(S + 0.5\omega_0 - j0.866\omega_0)(S + 0.5\omega_0 + j0.866\omega_0)}{(S + 0.25\omega_0 - j0.968\omega_0)(S + 0.25\omega_0 + j0.968\omega_0)} \tag{22-143}$$

In both Eqs. 22-142 and 22-143, ω_0 has the value 2000π. The two zeros in the numerator of Eq. 22-143 are complex conjugates of each other having negative real parts and thus are in the left half of the S plane which denotes minimum phase. Remember a zero is the value of S which makes the expression zero. Similarly, the two poles in the denominator are also complex conjugates of each other having negative real parts and thus are also located in the left half of the S plane. In order to equalize this anomaly, the equalizer must have a transfer function, H_e such that

$$H_r H_e = 1 \tag{22-144}$$

The transfer function of the required equalizer clearly must be just the reciprocal of the transfer function of the anomaly. Therefore,

$$H_e = \frac{(S + 0.25\omega_0 - j0.968\omega_0)(S + 0.25\omega_0 + j0.968\omega_0)}{(S + 0.5\omega_0 - j0.866\omega_0)(S + 0.5\omega_0 + j0.866\omega_0)} \tag{22-145}$$

The transfer function required of the equalizer is physically realizable as its poles are also in the left half of the S plane. Furthermore this transfer function is also minimum phase as its zeros are also in the left half of the S plane.

Figs. 22-54 and 22-55 display the Bode plots for the resonant anomaly while Figs. 22-56 and 22-57 present those of the equalizer.

If one adds the two amplitude response curves, Figs. 22-54 and 22-56, together point by point the result is a flat line at 0 dB. Similarly, if one adds the

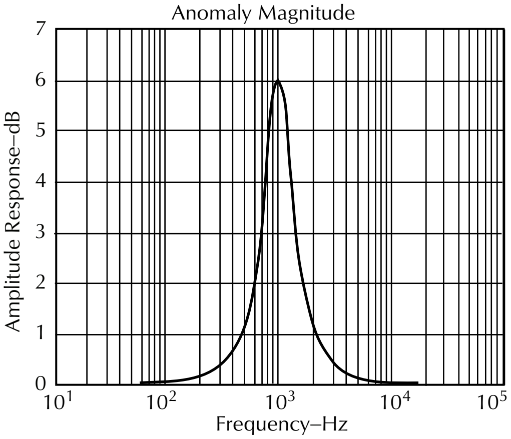


Figure 22-54. Amplitude response of resonance anomaly.

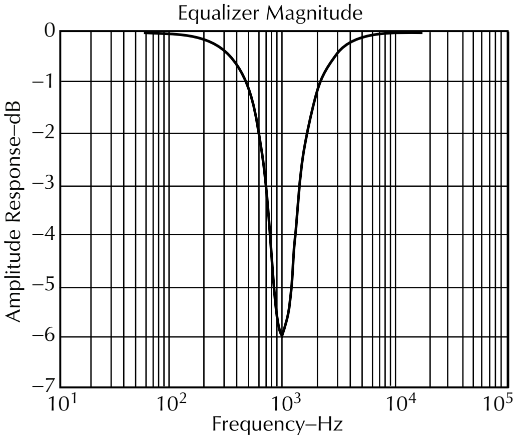


Figure 22-56. Amplitude response of corrective equalizer.

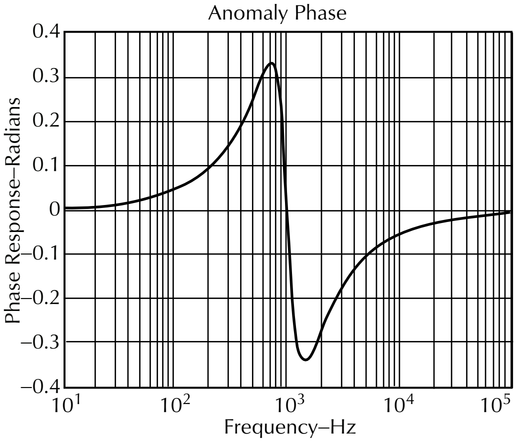


Figure 22-55. Phase response of resonance.

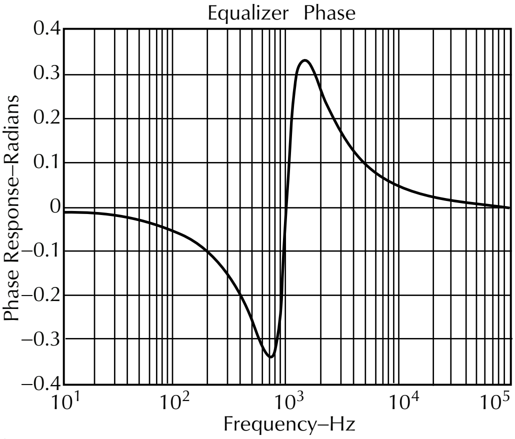


Figure 22-57. Required equalizer phase response.

phase response curves, Figs. 22-55 and 22-57, together point by point the result is a flat curve at 0 radian. In other words the effect of the resonance is exactly compensated for by the action of the equalizer. The resulting overall response with the equalizer in the signal chain will be that of a wide band loudspeaker which is devoid of any resonance as displayed in Fig. 22-58.

It should be remembered that the equalizer does not remove the resonance from the loudspeaker. The equalizer only *reduces* the amplitude and adjusts the phase of the drive signal in the vicinity of the resonance so as to negate the resonance's effects in the overall result.

The final curve displayed in Fig. 22-58 would also be that of an idealized loudspeaker with 3 dB down points at 20 Hz and 20 kHz and without any anomalies resonant or otherwise. Needless to say such a device, if available, would find a large market. The qualitative behavior at the high and low frequency ends of the response curve, however, reasonably matches many real loudspeakers except for the positions of the 3 dB down points. At the high frequency end, the behavior is that of a low pass filter while at the low frequency end, the behavior is that of a high pass filter. In those instances where this filter behavior is minimum phase, many practitioners employ equalizers to widen the bandwidth of narrow band loudspeakers. This process *increases* the amplitude and modifies the phase of the electrical drive signal at the frequency extremes that, indeed, will extend the range of response at the expense of requiring significantly larger displacements and power dissipation in the transducer. The larger displacements can easily lead to non-linear behavior and attendant distortion while the higher power can well produce thermal failure. Such practices, therefore, should be approached with extreme caution.

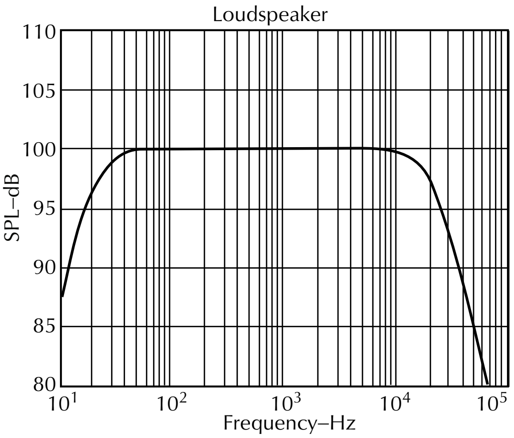


Figure 22-58. Overall response with equalizer in place.

Equalization is a powerful tool when intelligently applied. When so employed, it can make a system which is inherently good into an even better one. It cannot, however, convert a poor system into a good one. To quote an old southern expression, “You can’t make a silk purse from a sow’s ear!”

22.3.4 Equalization—Global or Local

Equalization is both the most often used as well as the most often abused signal processing function. The foregoing example in which a loudspeaker’s minimum phase type of anomaly was corrected is an example of global equalization. The correction that is invoked improves the quality of sound for all listeners regardless of where they might be located in the direct field of the loudspeaker. Global equalization of a loudspeaker is both a legitimate and desirable signal processing function. An accurate accomplishment of such an equalization often requires the services of a multi-band parametric equalizer. Such equalizers allow adjustments for filter frequency assignment as well as filter width and depth.

Local equalization, when it can be legitimately applied, improves sound quality at only one point while adversely affecting sound quality at all other points. In order to appreciate this we will explore a simple example where local equalization is employed to negate the effects of a single boundary reflection. Fig. 22-59 depicts the physical situation to be considered.

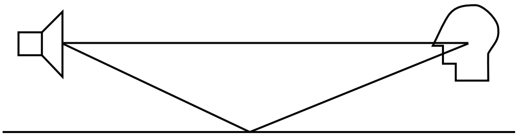


Figure 22-59. Direct sound accompanied by a single boundary reflection.

The listener in the figure only has the use of his left ear. He foolishly may have spent too many hours on the firing range without the benefit of hearing protection. Consider that the depicted coaxial loudspeaker is actually mounted in a properly sealed enclosure and that its crossover is seamless. The distances are such that the attenuated reflected sound arrives two milliseconds after the direct sound having undergone a broadband attenuation of 3 dB. We will be very generous in describing the loudspeaker response by giving it a second order Butterworth high pass and low pass characteristic with –3 dB points at 20 Hz and 20 kHz. This is much easier said than done! Upon letting $t = 0$ coincide with the arrival of the direct sound, the transfer function describing the direct sound aside from a scale factor can be written as

$$H = \frac{10^6 S^2 \omega_0^{\ 2}}{(S^2 + \sqrt{2} S\omega_0 + \omega_0^{\ 2})(S^2 + 10^3 \sqrt{2} S\omega_0 + 10^6 \omega_0^{\ 2})} \tag{22-146}$$

where,
f_0 is 20 Hz,
ω_0 is $2\pi f_0$.

The amplitude and phase responses associated with the direct sound appear in Fig. 22-60 and Fig. 22-61.

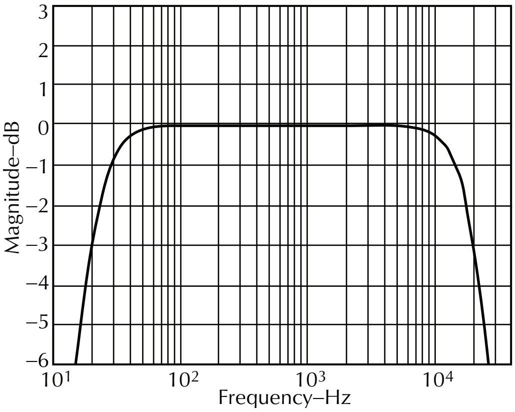


Figure 22-60. Direct sound amplitude response.

The reflection has an amplitude that is 3 dB down at all frequencies and is delayed in time by 2 ms relative to the direct sound and thus is described by the transfer function

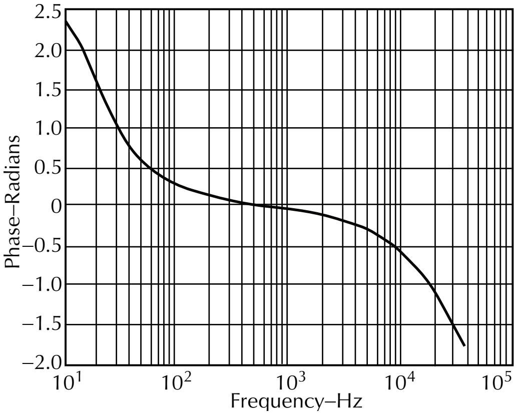


Figure 22-61. Direct sound phase response.

$$H_R = \frac{1}{\sqrt{2}}e^{-(0.002s)S} \tag{22-147}$$

The transfer function describing the combination of the direct and reflected sound is then

$$H_C = \left(\frac{10^6 S^2 \omega_0^{\,2}}{(S^2 + \sqrt{2}S\omega_0 + \omega_0^{\,2})(S^2 + 10^3\sqrt{2}S\omega_0 + 10^6\omega_0^{\,2})}\right) \times \left(1 + \frac{1}{\sqrt{2}}e^{-(0.002s)S}\right) \tag{22-148}$$

Aside from a scale factor, Eq. 22-148 describes the linear combination of the direct sound plus an attenuated delayed reflection at the listener's ear. One would expect that the spectrum of this signal would exhibit comb filter effects. This is indeed the case as exhibited in both the amplitude and phase behavior of this composite signal displayed in Figs. 22-62 and 22-63. Note that the displayed frequency range has been restricted to 4000 Hz because of the narrow spacing between interference notches.

The question at this point amounts to determining if it is possible to negate the effects of the reflection through some equalization technique. Mathematically, this amounts to inquiring what steps must be taken to reduce Eq. 22-148 to Eq. 22-146. From a mathematical point of view this can be done through the multiplication of Eq. 22-148 by the factor

$$\frac{1}{1 + \frac{1}{\sqrt{2}}e^{-(0.002s)S}}$$

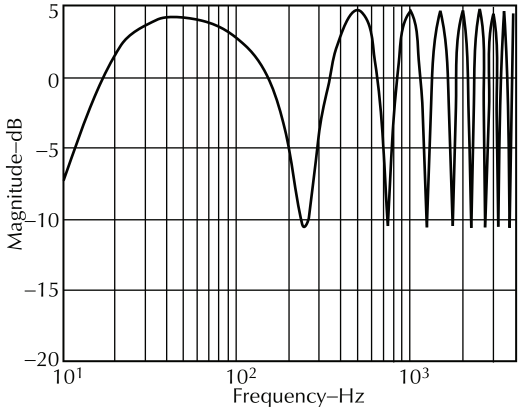


Figure 22-62. Amplitude response of direct plus reflected sound.

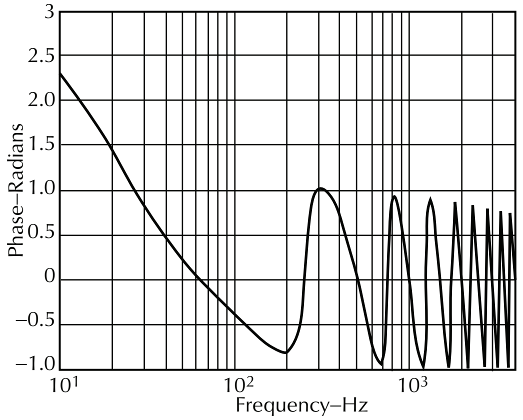


Figure 22-63. Phase response of both direct and reflected sound.

The transfer function of the equalizer required to perform this function thus must be

$$H_E = \frac{1}{1 + \frac{1}{\sqrt{2}}e^{-(0.002s)S}} \tag{22-149}$$

The question now becomes one of whether it is physically possible to construct such an equalizer. Remember, that in order to have a physically realizable stable system, the poles for the transfer function must have negative real parts. In examining Eq. 22-149 it is found that the numerator is a constant and hence there are no zeros associated with the transfer function. The poles of the transfer function are located at values of S for which the denominator becomes zero. These values of S are those for which

$$S = [-500\ln(\sqrt{2}) \pm jn500\pi] \tag{22-150}$$

where,

$n = 1, 3, 5, 7, \ldots$ all odd integers.

In principle this equalizer is physically realizable as all of the poles have negative real parts and thus lie in the left half of the complex plane. The poles are complex and considering all possible frequencies, are infinite in number. In practice one need consider only those that fall in the pass band of the loudspeaker. Even with this restriction there are at least 80 poles that must be considered. The amplitude and phase responses of the required equalizer are presented in Figs. 22-64 and 22-65.

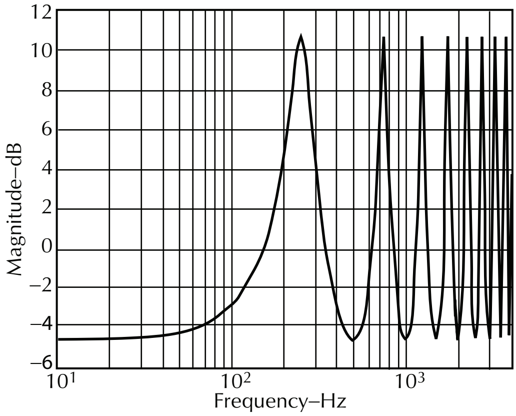


Figure 22-64. Equalizer amplitude response.

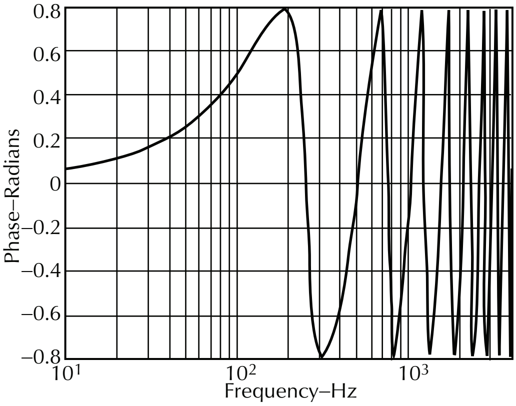


Figure 22-65. Equalizer phase response.

The good news is that if an equalizer having a transfer function described by H_E is inserted in the electronic chain then the one-eared observer will experience only what would have been just the direct sound from the loudspeaker. This is accomplished by distorting the drive signal to the loudspeaker in such a way as to completely negate the effects of the reflection. Notice that this requires a drive signal boost greater than 10 dB at certain frequencies accompanied by a cut greater than 4 db at other frequencies. The boost of over 10 dB could easily lead to headroom difficulties. Now for the bad news. Even though the one-eared observer will be well pleased, what about those observers exposed to only the direct sound? The other observers will be exposed to a sound field characterized only by the product of the loudspeaker transfer function with the transfer function of the equalizer. These observers will experience a frequency response described in Figs. 22-66 and 22-67.

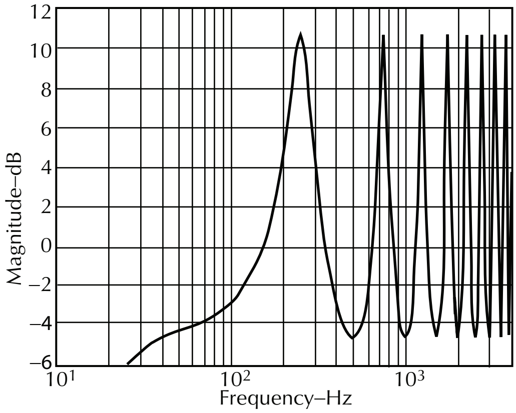


Figure 22-66. Equalized amplitude response for observers not exposed to the reflection.

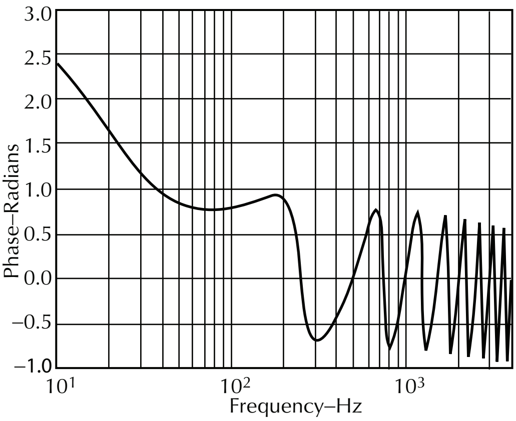


Figure 22-67. Equalized phase response for observers not exposed to the reflection.

In summary, global equalization should always be applied to the direct field of the loudspeaker as this improves sound for all observers. Local equalization may indeed be possible but its application may improve the sound at a selected point while worsening the sound experienced at other points. As a general rule, equalization should be applied with care as one may well experience headroom difficulties leading to distortion and perhaps loudspeaker thermal failure.

All of the foregoing analysis has dealt with analog systems in which time is a continuous variable and the signals involved were analog variables such as continuous time dependent voltages, currents, acoustic pressures, etc. We turn next to digital systems where time is measured in discrete intervals and the dependent variables are a sequence of binary encoded values such as those issuing from an ADC. System theory in this instance is based not on the Laplace transform but rather on a related discrete variable transform known as the *Z* transform. It is important not to confuse the employment of the *Z* symbol in connection with discrete systems with the employment of the same symbol *Z* in describing electrical impedance.

22.4 Digital Systems and the *Z* Transform

We will begin our study of digital system theory by first examining an analog system that also can be modeled digitally in an almost intuitive manner. In connection with an analysis of the ensuing analog and digital models Table 22-6 will prove to be of great value.

Table 22-6. System Analysis Functions in Both Analog and Digital Domains

Items	Analog Domain	Digital Domain
Time	t	nT
Dependent function	$f(t)$, $g(t)$, etc.	$f(nT)$, $g(nT)$, etc.
Unit impulse	$\delta(t)$	1
Complex plane variable	S	$Z = e^{ST}$
Transform	$F(S) = \int_0^\infty f(t)e^{-St}dt$	$F(Z) = \sum_{n=0}^{n=\infty} f(nT)Z^{-n}$

The analog system consists simply of a two-input summer and a length of transmission line such as a section of co-axial cable (assumed to be lossless and properly terminated). The signal transit time along the cable from input to termination is taken to be a fixed time denoted by *T*. This same time interval *T* will also be taken as the sampling interval in the corresponding digital system. A single line drawing of the analog system appears in Fig. 22-68.

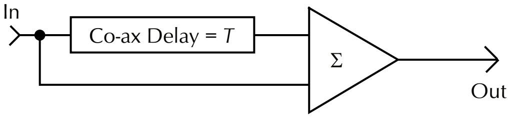


Figure 22-68. This is a simple analog system that is to serve as a model for a corresponding digital system.

In performing system analysis of the analog system of Fig. 22-68 we consider how the system behaves under impulsive excitation so that the input signal is taken to be $\delta(t)$. The output signal is then simply the sum of the input with a version of the input that has been delayed by a time *T*. Upon denoting the input signal by $f(t)$ and the output signal by $g(t)$ we can write

$$\begin{aligned} f(t) &= \delta(t) \\ g(t) &= \delta(t) + \delta(t - T) \end{aligned} \qquad (22\text{-}151)$$

The next step is to take the Laplace transform of the input and the Laplace transform of the output. The Laplace transform of the unit impulse is just 1 and the Laplace transform of the time shifted unit impulse is just e^{-ST} times the Laplace transform of the unit impulse itself so that

$$\begin{aligned} F(S) &= 1 \\ G(S) &= 1 + e^{-ST} \end{aligned} \qquad (22\text{-}152)$$

The transfer function of the system is the quotient of the output Laplace transform by the input Laplace transform and is thus

$$\begin{aligned} H(S) &= \frac{G(S)}{F(S)} \\ &= (1 + e^{-ST}) \end{aligned} \qquad (22\text{-}153)$$

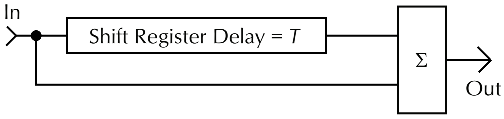


Figure 22-69. The corresponding digital system.

The corresponding digital system single line drawing is represented by Fig. 22-69.

The digital system of Fig. 22-69 has its input provided by the data register at the output of the ADC. The input is described by the numerical sequence $\{f(nT)\}$ where n can range over all integer values from zero to infinity. Similarly the output is described by the numerical sequence $\{g(nT)\}$. If no signals have been applied prior to $n = 0$, the contents of the shift register will be 0 as will be the contents of the ADC's data register. For $n = 0$, the data register at the output of the ADC assumes a value of 1 while the contents of the shift register remain at 0. When n takes on the value of 1 the contents of the data register become zero and remain there as no other signals are considered to be applied to the system. On the other hand, when n takes on the value of 1, the shift register acquires the value 1 that was formerly contained in the ADC's register. The shift register always lags the data register by one sample thus introducing a delay of T just as did the co-axial cable in the analog system. The summer presents at its output the combined contents of both the data register and the shift register. This behavior is summarized in Table 22-7.

Table 22-7. Digital Example Data Table

n	$f(nT)$	$g(nT)$
0	1	1
1	0	1
2	0	0

For the digital system under consideration the exciting signal is considered to be a single unit impulse applied at $n = 0$ hence all entries in the table beyond $n = 1$ will be 0. The transfer function of the system is now calculated by taking the Z transform of the output and then dividing this transform by the Z transform of the input. By following the description of the Z transform listed in Table 22-6 these transforms are found to be

$$\begin{aligned} G(Z) &= \sum_{n=0}^{n=\infty} g(nT)Z^{-n} \\ &= \sum_{n=0}^{n=1} g(nT)Z^{-n} \\ &= (1+Z^{-1}) \\ F(Z) &= \sum_{n=0}^{n=\infty} f(nT)Z^{-n} \\ &= \sum_{n=0}^{n=1} f(nT)Z^{-n} \\ &= 1 \end{aligned} \quad (22\text{-}154)$$

In writing the results of Eq. 22-154 it was recognized that $Z^{-0} = 1$. The transfer function describing the operation of this digital system is thus

$$\begin{aligned} H(Z) &= \frac{G(Z)}{F(Z)} \\ &= (1+Z^{-1}) \end{aligned} \quad (22\text{-}155)$$

When one substitutes the value of Z as listed in Table 22-6 into Eq. 22-155 it is found that the transfer function of the digital system has a mathematical form identical to that of the analog system from which it was modeled. This is an exceptional case in that regard as will be learned when other digital systems are examined. Many readers will recognize that both the analog as well as the digital system considered thus far constitute a comb filter as they both simply add a given signal to a delayed version of itself. There is a significant difference, however, because in the analog world the delay interval can take on any value whereas in the digital world it is restricted to integral multiples of the sampling interval. This will turn out to have far reaching consequences. The digital system of this example is a special case of a digital filter type that is referred to generically as a finite impulse response filter or FIR. It is called a finite impulse response filter because its response to a unit impulse exists for a finite interval of time. This is in contrast to another type of digital filter known as the infinite impulse response filter or IIR. The impulse response of a stable IIR in principle endures for all time though all the while diminishing as time increases. The structures of IIRs differ from those of FIRs in that the IIRs employ feedback from output to input whereas the FIRs employ only feed-forward techniques.

We return now to further analysis of our simple analog and digital filters of the present example. In the analog case, the complex frequency variable is S and in the steady state S takes on the value $j\omega = j2\pi f$ where the signal frequency variable in principle can range up to an infinite value. In the digital case, the complex frequency variable is Z where $Z = e^{ST}$ and T is the reciprocal of the sampling frequency. In the steady state S is again $j2\pi f$ but the operating frequency can range only up to one-half of the sampling frequency. Beyond this limit aliasing rears its ugly head. The maximum operating frequency in a practical case is always held to less than this limit prior to the digital sampling process. If we let f_m represent the operating frequency maximum at the aliasing limit and f represent the signal frequency variable then it is possible to write

$$\begin{aligned} T &= \frac{1}{2f_m} \\ Z &= e^{j\pi\frac{f}{f_m}} \end{aligned} \quad (22\text{-}156)$$

The absolute magnitude of the exponential function representing Z is unity independent of the operating frequency and hence the imaginary frequency axis of the S plane between $-jf_m$ and $+jf_m$ is bent into a circle of unit radius when displayed or mapped into the Z plane. The interior of this circle contains the entire contents of the left half of the S plane between these frequency limits. This mapping is displayed in Fig. 22-70.

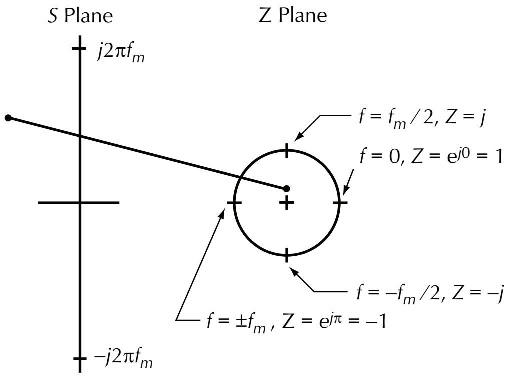


Figure 22-70. S plane to Z plane mapping.

In the steady state description of frequency all of the points of interest lie on the unit circle as $S = j2\pi f$. In doing pole-zero analysis, $S = \sigma + j2\pi f$ and such points are described in the Z plane in terms of polar coordinates r, θ. These coordinates relate to those in the S plane as expressed in

$$\begin{aligned} r &= e^{\sigma T} \\ \theta &= \pi\frac{f}{f_m} \end{aligned} \qquad (22\text{-}157)$$

A typical point is indicated as a black dot in the S plane. The coordinates of this particular point are $S = -(\pi/2T) + j(\pi/2T)$. The frequency value of this point in the S plane corresponds to $f_m/2$ so it will appear in the Z plane with $\theta = \pi/2$. The radial distance from the origin will be $r = e^{-\pi/2}$. This corresponds to a radial distance of about 0.2. Therefore the Z plane point corresponding to the given S plane point appears at a position immediately above the origin of the circle at a distance of 0.2. The angle of this location is the angle between the radial line drawn to the point and the horizontal or real axis and is the required $\pi/2$. The given S plane point is thus mapped into the Z plane.

The comb filter of our analog and digital example is a crude form of low pass filter. The frequency response in the digital case is obtained by substituting $S = j2\pi f$ into Eq. 22-155. The amplitude response is given by the magnitude of the resulting complex expression while the phase response is given by the angle of the complex expression. After the substitution for S has been made and Euler's Theorem is applied to the result, the expression for the transfer function becomes

$$H = 1 + \cos\left(\pi\frac{f}{f_m}\right) - j\sin\left(\pi\frac{f}{f_m}\right)$$

Upon calculating the magnitude of this complex expression followed by some simplification, the amplitude response appears as

$$|H| = 2\cos\left(\frac{\pi f}{2f_m}\right) \qquad (22\text{-}158)$$

The arctangent of the ratio of the imaginary to the real part of the transfer function yields the phase response. This expression can also be simplified to yield

$$\varphi = -\frac{\pi f}{2f_m} \qquad (22\text{-}159)$$

This last result is very significant as it indicates that FIR filters are capable of producing exactly linear phase responses.

A general FIR is constructed from a chain of shift registers or a dedicated read-write memory space along with a collection of multipliers and a summer or accumulator as depicted in Fig. 22-71. The transfer function for this filter is given by

$$H(Z) = \sum_{i=0}^{i=N-1} c_i(Z^{-1})^i \qquad (22\text{-}160)$$

In Eq. 22-160 N is the number of taps possessed by the filter, the number of independent multiplier coefficients available for shaping the filter's behavior, and the number of samples over which the filter's impulse response endures. If N were equal to three then the transfer function would be $H(Z) = c_0 + c_1Z^{-1} + c_2Z^{-2}$ and the impulse response of this filter would endure over only 3 samples. FIR digital filters are often called upon to have amplitude responses that mimic those of standard analog filters such as Butterworth, Chebyshev, etc.

One technique for accomplishing this is to adjust the filter's multiplier coefficients in such a manner that the filter's impulse response is a discrete time version of the model analog filter's continuous time impulse response. Any reasonable degree of accuracy in fashioning the impulse response in this manner can require a structure having a large number of taps.

As an example we will take an analog first order low pass filter as a structure whose amplitude

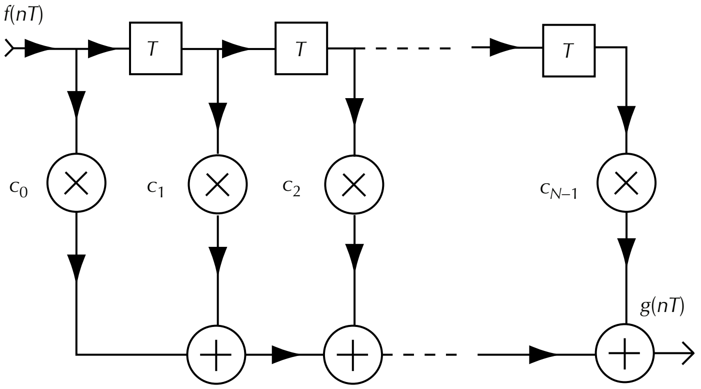


Figure 22-71. Structure of an *N* tap FIR.

response is to be mimicked by an FIR filter. It was learned in continuous time system theory that the impulse response of such a filter has a time behavior of the form $e^{-(t/\tau)}$ where τ is the time constant of the analog low pass filter. The impulse response is matched between the continuous and discrete time domains by requiring $e^{-(t/\tau)} = e^{-(nT/\tau)}$ where *n* begins with zero and takes on the sequence of positive integers. The tap multiplier coefficients for the FIR filter are then given by

$$c_i = e^{-\frac{iT}{\tau}} \tag{22-161}$$

In Eq. 22-161 *i* begins with zero and takes on the sequence of positive integers up to the limit set by the number of taps employed by the filter. Eq. 22-161 is now substituted into Eq. 22-160 to obtain an expression for calculating the desired transfer function.

$$H(Z) = \sum_{i=0}^{i=N-1} \left(e^{-\frac{T}{\tau}} Z^{-1}\right)^i \tag{22-162}$$

What is desired at this point is a closed form expression for the transfer function in which the number of taps, *N*, appears explicitly as a variable. This will allow a trial and error method for choosing the least number of taps required to obtain the desired degree of accuracy of performance. Denoting the total expression within the parentheses of Eq. 22-162 by the symbol *x* we will make use of a well-known property of power series. When the absolute magnitude of *x* is equal to or less than one then

$$\sum_{i=0}^{i=\infty} x^i = \frac{1}{1-x} \tag{22-163}$$

Upon applying this, Eq. 22-162 can be written as

$$\begin{aligned} H(Z) &= \sum_{i=0}^{i=N-1} x^i \\ &= \left(\sum_{i=0}^{i=\infty} x^i - \sum_{i=N}^{i=\infty} x^i\right) \end{aligned} \tag{22-164}$$

Further exploration of Eq. 22-164 yields

$$\begin{aligned} H(Z) &= \frac{1}{1-x} - \frac{x^N}{1-x} \\ &= \frac{1-x^N}{1-x} \end{aligned} \tag{22-165}$$

At this point we replace *x* by the expression in the parentheses of Eq. 22-162 for which it stands to obtain

$$H(Z) = \frac{1-\left(e^{-\frac{T}{\tau}} Z^{-1}\right)^N}{1-e^{-\frac{T}{\tau}} Z^{-1}} \tag{22-166}$$

Finally, we recall that $Z^{-1} = e^{-ST}$ and in obtaining the frequency response we must make the substitution $S = j2\pi f$. The amplitude response of the filter is obtained by taking the absolute magnitude of Eq. 22-166 after these substitutions are made. The procedure at this point is then to take a trial value for *N*, calculate the frequency response, and compare its plot with that of the model analog filter. The procedure is then to find the least value of *N* that yields satisfactory results. The processing time in the FIR, and hence its latency, is directly proportional to *N*, so a premium is placed on small *N*.

When the above procedure is applied in generating an FIR that mimics a first order analog low pass whose cut off is at 500 Hz, it is found that good results are obtained for $N = 100$ as displayed in Fig. 22-72.

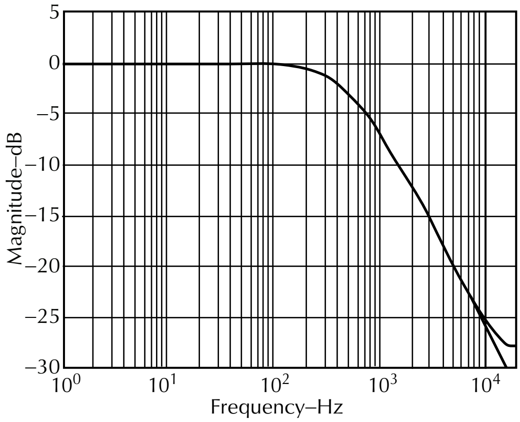


Figure 22-72. Analog to FIR digital amplitude response comparison.

In constructing Fig. 22-72 the transfer function of the FIR has been normalized to match that of the analog filter at zero frequency. A departure between the analog and digital frequency responses occurs only in the vicinity of the Nyquist limit that in this instance is at 20 kHz as the sampling frequency was taken to be 40 kHz. The design technique employed in the foregoing is not unique. It has been pursued here as it is perhaps the most intuitive. There are even more refined design techniques available for designing both FIR and IIR filters. The author has found those available in Matlab to be most valuable.

22.4.1 Recursive or IIR Filters

In the continuous time domain the laws governing system operations are stated in the form of linear ordinary differential equations. The corresponding structure in the discrete time domain is called a difference equation. In becoming familiar with this we will again take an analog first order low pass filter as our model and implement it now by means of an IIR structure. The analog filter is governed by a first order differential equation suggesting that the starting point for the corresponding IIR filter will be a first order difference equation. If $\{f(nT)\}$ is the input signal sequence and $\{g(nT)\}$ is the output signal sequence, then the governing equation is written as follows with a and b playing the roles of numerical constants.

$$g(nT) = af(nT) - bg([n-1]T) \qquad (22\text{-}167)$$

This equation is implemented by the structure of Fig. 22-73.

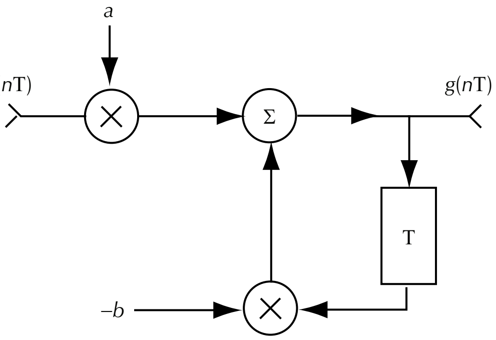


Figure 22-73. Circuit implementation of first order difference equation.

The transfer function of the circuit of Fig. 22-73 is found by taking the Z transform of each term in Eq. 22-167 and solving for the ratio of the transform of the output divided by that of the input. In doing so it should be recalled that the Z transform of a unit delay term is the transform of the undelayed term multiplied by Z^{-1} therefore

$$\begin{aligned} G(Z) &= aF(Z) - bZ^{-1}G(Z) \\ H(Z) &= \frac{G(Z)}{F(Z)} \qquad (22\text{-}168) \\ &= \frac{a}{1 + bZ^{-1}} \end{aligned}$$

The transfer function of Eq. 22-168 can be made to conform to that of a first order low pass filter through a suitable choice for the constants a and b. Instead of matching impulse responses as was done in the FIR example (although that could well be done here) a different technique will be introduced. In this technique the pole location of the model analog filter will be mapped into the corresponding location in the Z plane. The transfer function for the model analog filter can be written as $1/(\tau S + 1)$. The pole for this transfer function is located at $S = -(1/\tau)$ where τ is the time constant of the filter either RC or L/R as the case may be. In the analog case this pole is located on the negative real axis at the point $\sigma = -(1/\tau)$. One now substitutes $Z = e\sigma^{T}$ into $H(Z)$ and locates the pole by requiring the denominator to be zero. One then solves the resulting equation for the constant b.

$$1 + be^{\frac{T}{\tau}} = 0$$

$$b = -e^{-\frac{T}{\tau}} \tag{22-169}$$

The transfer function can now be written as

$$\frac{a}{1 - e^{-\frac{T}{\tau}} e^{-j\pi\frac{f}{f_{max}}}} \tag{22-170}$$

Finally, in the analog case the amplitude response is unity at zero frequency. When this is also required of the digital filter, the final form of the transfer function becomes

$$H(Z) = \frac{1 - e^{-\frac{T}{\tau}}}{1 - e^{-\frac{T}{\tau}} e^{-j\pi\frac{f}{f_{max}}}} \tag{22-171}$$

A comparison between the amplitude behavior of this digital filter and its analog counterpart is presented in Fig. 22-74. The sampling frequency in this instance is again 40 kHz. As a result the amplitude performance of the digital filter does not match that of the analog as the frequency approaches the aliasing limit of 20 kHz. The performance can be improved in this region through the employment of a considerably higher sampling frequency.

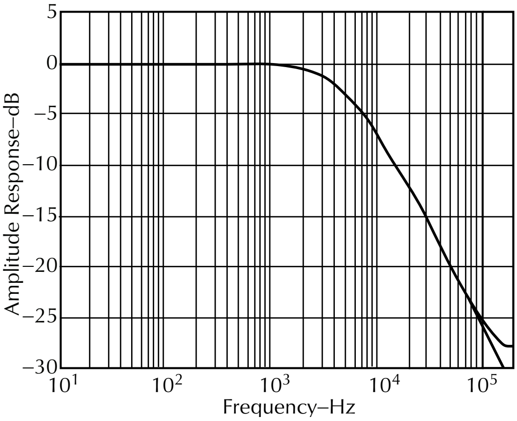


Figure 22-74. IIR and analog amplitude response comparison.

Fig. 22-74 contains only part of the story. One needs to also compare the phase responses of the two filters. This comparison is presented in Fig. 22-75.

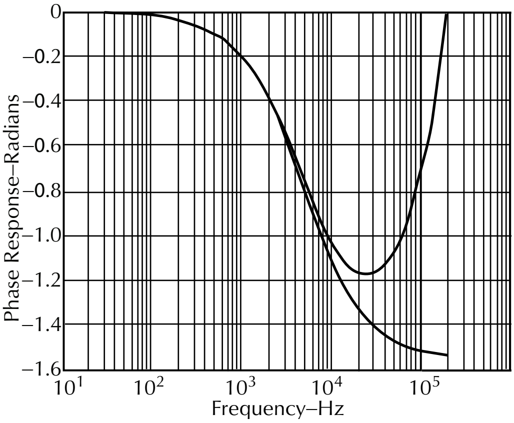


Figure 22-75. IIR and analog phase response comparison.

The well-known minimum phase behavior of the analog filter is displayed in the lower curve of Fig. 22-75. The digital filter tracks this at low frequencies but diverges markedly above about 2 kHz. A higher sampling frequency will bring about some improvement but nevertheless non-minimum phase behavior will still remain. The aberrant phase behavior could be corrected through the employment of suitable all pass digital filters. This, however, would require considerably more hardware as well as design effort. There exist design algorithms that evolve IIR filters that simultaneously match the corresponding analog versions in both amplitude as well as phase responses. These algorithms are based on the employment of second order filter sections.

A second order section can produce a first order section through the appropriate choice of constants and filters of any order may be obtained by cascading suitable numbers of second order sections. The second order difference equation that governs the structure of a second order section can be written as

$$g(nT) = a_0 f(nT) + a_1 f([n-1]T) + a_2 f([n-2]T) - b_1 g([n-1]T) - b_2([n-2]T) \tag{22-172}$$

Upon taking the *Z* transform of Eq. 22-172 one can solve for the transfer function to obtain

$$\begin{aligned} H(Z) &= \frac{G(Z)}{F(Z)} \\ &= \frac{a_0 + a_1 Z^{-1} + a_2 Z^{-2}}{1 + b_1 Z^{-1} + b_2 Z^{-2}} \end{aligned} \tag{22-173}$$

This general transfer function for the second order section is called a biquad as it is quadratic in *Z* in both the numerator and the denominator. The general biquad is implemented by the arrangement shown in Fig. 22-76.

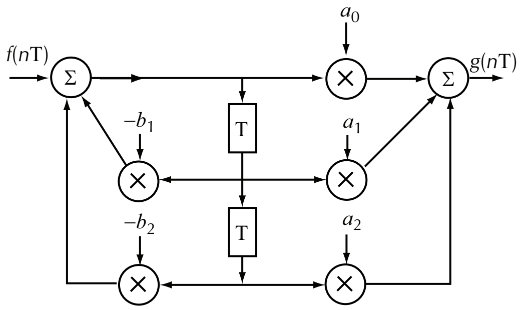


Figure 22-76. Hardware implementation of the general biquad.

We are now in position to correct the aberrant phase behavior encountered earlier with regard to the first order low pass filter. The procedure is to map the pole or poles of the analog prototype exactly in the Z plane. As the low pass was first order with just a single pole then the coefficients a_2 and b_2 in Eq. 22-173 are set equal to zero. The phase behavior is corrected by placing a zero in the numerator that is not present in the analog prototype. This zero is located at the aliasing limit so as to have a minimum effect on the amplitude response. Recall that the amplitude response was correct in the original simple design. This step will require that a_0 and a_1 be equal. Finally the common value of these coefficients is adjusted to normalize the amplitude response of the digital filter to equal that of the analog prototype in the filter pass band. Finally in order to make the filter more in accord with actual practice we will now employ a sampling rate of 48 kHz. The transfer function for this digital filter now becomes

$$H(Z) = \frac{0.031699(1+Z^{-1})}{1-0.93660Z^{-1}} \qquad (22\text{-}174)$$

The comparative performance of this improved digital filter is presented in Figs. 22-77 and 22-78.

The phase behavior of this improved design is almost an exact replica of the analog prototype while the amplitude error is at most about 1 dB at the frequency extreme of 20 kHz. This filter then furnishes the required minimum phase behavior while having only a minor amplitude error at the frequency extreme. One final example that employs the quadratic structure of the biquad is that of a bandpass filter. We will consider a unity gain bandpass filter having a bandwidth of one octave that is

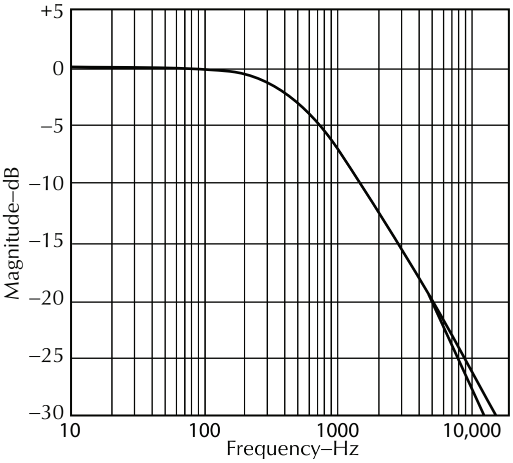


Figure 22-77. Analog and digital amplitude comparison for an improved IIR filter.

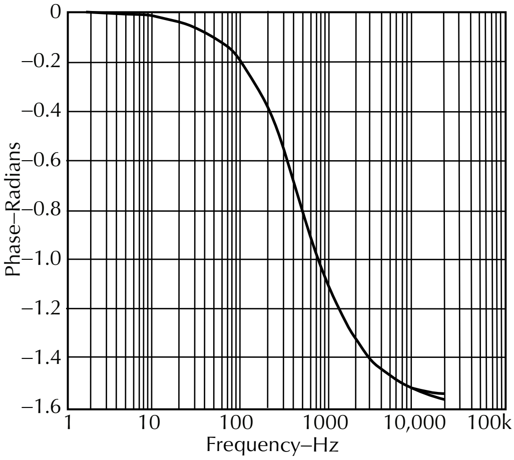


Figure 22-78. Analog and digital phase comparison for an improved IIR filter.

centered on 500 Hz. The transfer function of such a filter with a 48 kHz sampling rate is given by

$$H(Z) = \frac{0.02262(1-Z^{-2})}{1-1.9506Z^{-1}+0.95476Z^{-2}} \qquad (22\text{-}175)$$

The complex Z plane poles of this filter have locations corresponding to those of the analog prototype. There are two zeros however. One of these is located at ω equal to zero as in the analog case while the second one that corrects the phase response is located at the aliasing limit. The comparative performance of the IIR and its analog prototype is displayed in Figs. 22-79 and 22-80.

An examination of the figures reveals that the amplitude error is at most about 1 dB at 20 kHz while the phase curves overlay each other nearly exactly.

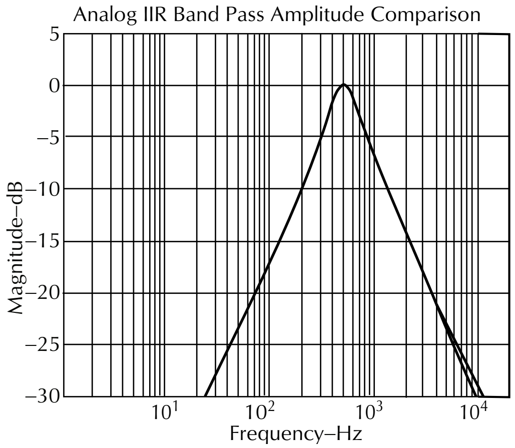


Figure 22-79. Analog and digital filter amplitude comparison for an octave bandpass.

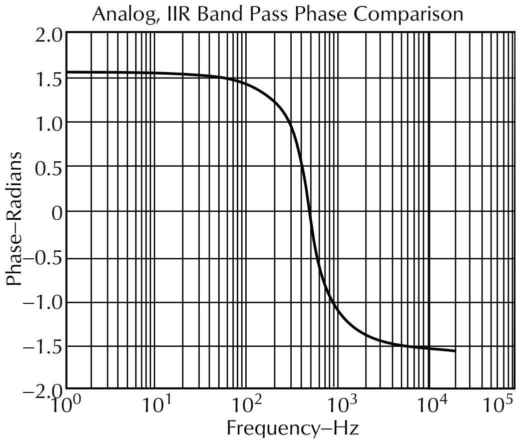


Figure 22-80. Analog and digital filter phase comparison for an octave bandpass.

All of the time honored analog filters have accurate IIR digital versions that can be implemented through the employment of the appropriate number of second order sections. The Linkwitz-Riley fourth order filter, for example, would require a cascade of two such sections with each being designed to conform to a second order Butterworth. The IIR filters employ feedback and hence can be unstable if not properly designed. Stable filter structures in the analog world require that all poles lie in the left half of the S plane. The corresponding criterion in the Z plane is that all poles are contained within the unit circle. This is not a concern with FIR filters as these filters do not employ feedback.

22.4.2 Linear Phase Filters

Certain FIR filters can be easily designed to possess a linear phase shift characteristic. This is a property that can only be obtained with difficulty through the employment of analog filters. Linear phase shift filters have no phase distortion in that they feature a constant group delay and are highly desirable in many applications. Linear phase shift requires that $\varphi = -\alpha\omega \pm \beta$. The group delay is the negative of the derivative of this expression with respect to ω and thus is equal to α independent of the constant β. When β is equal to zero, the impulse response of the filter in the Fourier sense is symmetric. When β is $\pi/2$, the impulse response is anti-symmetric.

Most filter designs start with a requirement for a particular amplitude response versus frequency. In the case of digital filters, if one has a mathematical description of the desired amplitude response ranging over both positive and negative frequencies it is possible to take the inverse discrete Fourier transform of the amplitude response and obtain the discrete time impulse response of the filter. When considering both positive and negative frequencies, the description of the amplitude response will be a real and even mathematical function.

The impulse response sample sequence will then be both real and even and thus symmetric about the origin. Recall from the classical uncertainty principle that when the amplitude response exists only over a relatively small range of the frequency axis as is true in the case of the audio spectrum, then the impulse response must exist over quite a large range of the discrete time axis. This means that the impulse response sample sequence $\{h(nT_S)\}$ is symmetric about $n = 0$ and has significant values out to large values of n in both the positive and negative directions. In our earlier simple introduction to the FIR filter we learned that the various values of the impulse response sequence determined the tap weights of the FIR filter. In order to match exactly the prescribed amplitude response function, the filter would have to accommodate the entire impulse response sequence and would be too long to be practical. It is necessary to truncate the impulse response sequence. This truncation introduces differences or errors between the desired response and that which is actually achieved. These errors will be larger for short FIRs and vice versa. We will now illustrate the manner in which a causal FIR accommodates a truncated impulse response. In order to keep the numbers involved simple the example will deal with an FIR that is too short for use in practice but will clearly illustrate what is involved. The FIR of the example has ten delay sections and eleven tap weights and thus $N = 11$. The tap weights or multiplier values are c_i with i ranging from 0 to 10. In Table 22-8 the tap weights are positioned above their respective values in terms of the impulse response sequence values.

Table 22-8. Time Shift of Truncated Impulse Response

c_0	c_1	c_2	c_3	c_4	c_5
$h(-5)$	$h(-4)$	$h(-3)$	$h(-2)$	$h(-1)$	$h(0)$
c_6	c_7	c_8	c_9	c_{10}	
$h(1)$	$h(2)$	$h(3)$	$h(4)$	$h(5)$	

The filter must accommodate the entire truncated impulse response in order to approach the desired amplitude response. In doing so it is necessary to time shift the impulse response by 5 samples. It is this shift in the time domain that brings about the linear phase response in the frequency domain. The final point to be dealt with is the question of the errors that were introduced because the impulse response was truncated. There exists a technique from approximation theory called the Remez exchange algorithm that can minimize and distribute the errors in a meaningful and useful manner. This is accomplished through small modifications to the tap weights.

All of the above steps in FIR design have been incorporated in computer based mathematics programs. A notable such program is available in Matlab© where it is described as the Remez, Parks-McClellan optimal equiripple FIR filter design process. As an example, this process was applied toward the design of a high pass filter with a pass band edge or cut off at 2 kHz. The sampling rate employed was 48 kHz. With this sampling rate, the Nyquist frequency or f_{max} is 24 kHz. Excellent performance was desired in the filter. The design process was run through several iterations with the number of taps being the variable. The final design employed 151 taps. As a consequence, the impulse response undergoes a shift of $(N-1)/2$ or 75 samples. The delay time through the filter is thus 75 times the sampling period of 1/48,000 s or 1.5625 ms. The filter's performance is depicted in Fig. 22-81.

Actually there are ripples in the amplitude response in both the pass band and stop bands of the filter. These ripples are so small that they do not appear with the amplitude scale employed. The ripple amplitude is small as a result of the large number of taps employed. As a comparison, the filter was also designed while specifying just 21 taps. The performance in this case is displayed in Fig. 22-82.

In Fig. 22-82 the ripples are more apparent for two reasons. The ripples have larger amplitude as a result of the reduced number of taps and the vertical scale has been adjusted to more clearly display the ripple effect. Notice also the overall phase change is now much less as the signal is now delayed by only ten samples.

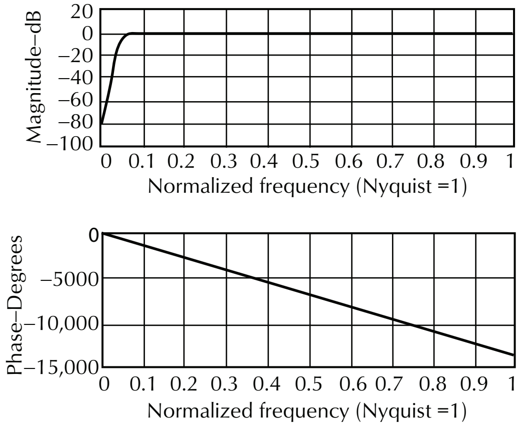


Figure 22-81. High pass FIR filter with 151 taps.

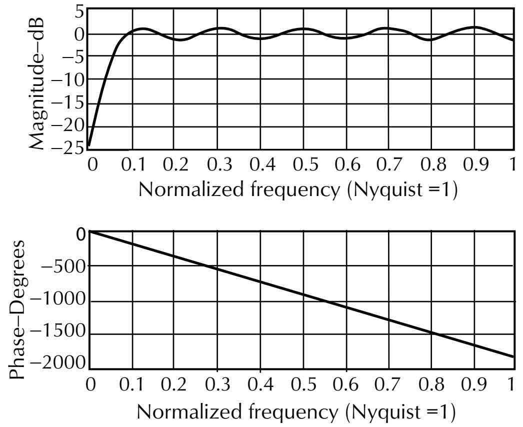


Figure 22-82. High pass FIR filter with 21 taps.

The hardware types are the same for both FIR and IIR filters. The differences appear only in the hardware configurations and in the amounts of hardware required to accomplish a given filter's task. Digital signal processing engines or DSPs contain collections of the necessary multipliers, summers, shift registers, memories, and controllers that may be configured to accomplish the various signal processing tasks in the discrete time domain.

22.5 Dynamics Processing

A dynamics signal processor sets the relationship that exists between the changes in the voltage levels at its output as compared with the changes in the voltage levels at its input. This relationship may in fact in some instances be a linear one. Even so, this does not mean that the voltage level at the output is the same as the voltage level at the input. Fig. 22-83 graphically displays three instances of linear processing.

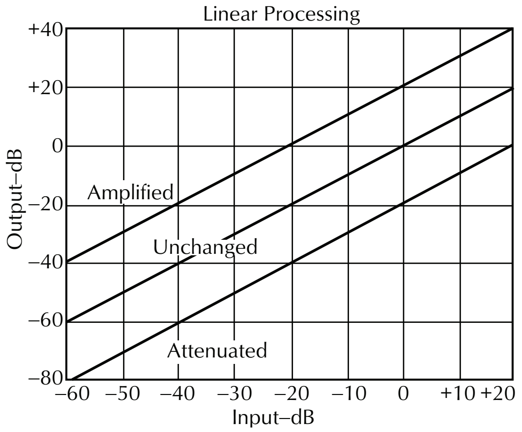


Figure 22-83. Three examples of linear processing.

In each of the three plots in Fig. 22-83 the processing is said to be linear as a change of 10 dB in input level causes a corresponding change of 10 dB in output level. Notice, however, that the output level in the upper curve is always 20 dB greater than that of the input indicating amplification by a factor of 10. In the middle curve the input level and output level are equal, indicative of unit amplification. In the lower curve, the output level is uniformly less than the input level by 20 dB thus indicating that the output is attenuated by a factor of 10 relative to the input. This behavior is expressed mathematically in the following fashion. Let V_o be the output voltage either amplitude or root mean square value. Similarly, let V_i be the corresponding input voltage value and let a be a dimensionless constant. In linear processing these quantities are related through

$$V_o = aV_i. \qquad (22\text{-}176)$$

The display in Fig. 22-83 deals with voltage levels rather than just input and output voltages per se. It is necessary then to divide both sides of Eq. 22-176 by a voltage reference value, take the logarithm to the base ten on both sides and multiply both sides by 20 dB to obtain

$$20\ \text{dB}\log\left[\frac{V_o}{V_r}\right] = 20\ \text{dB}\log[a] + 20\ \text{dB}\log\left[\frac{V_i}{V_r}\right] \qquad (22\text{-}177)$$

In writing Eq. 22-177 use has been made of the fact that the logarithm of a product is the sum of the logs of the multiplier and the multiplicand. When one assigns a the values of 10, 1, and 0.1, respectively, Eq. 22-177 generates the upper, middle, and lower plots of Fig. 22-83.

Non-linear dynamics processing is more interesting and is represented by the two distinctly different categories termed expansion and compression. Expansion occurs when a small change in input level produces a larger change in output level such as a 1 dB change in input level producing a 2 dB change in output level.

This is described as being 1 into 2 or 1:2. Similarly compression occurs when a large change in input level produces a smaller change in output level such as a 2 dB change in input level producing a 1 dB change in output level. This is described as being 2 into 1 or 2:1. Fig. 22-84 graphically displays both of these expansion and compression curves along with a unity amplification linear reference.

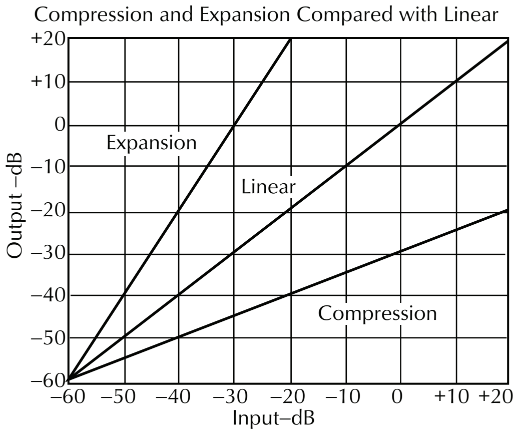


Figure 22-84. Expansion 1 into 2 and compression 2 into 1 as compared with linear.

The equation for the expansion curve can be written by inspection of the upper curve in Fig. 22-84.

$$\begin{aligned} 20\ \text{dB}\log\left[\frac{V_o}{V_i}\right] &= 2(20\ \text{dB})\log\left[\frac{V_i}{V_r}\right] \\ &= 20\ \text{dB}\log\left[\frac{V_i^2}{V_r^2}\right] \end{aligned} \qquad (22\text{-}178)$$

If one solves Eq. 22-178 for V_o in terms of V_i it will be found that V_o is proportional to V_i^2. A similar analysis applied to the compression curve shows that for compression V_o is proportional to $V_i^{1/2}$. In general then, if n is the exponent applied to V_i, expansion occurs for $n > 1$ while compression occurs for $n < 1$. If one desires an expansion of 1 dB into 4 dB then n must have the value 4. If one desires a compression of 4 dB into 1 dB then n must have the value 1⁄4. Both expansion and compression are non-linear as a power law relates output to input. If one applies expansion with a quite large value for n such as $n = 10$ or more the process is termed gating. Similarly if one applies compression with $n = 1/10$ or 1⁄20 the process is called limiting. Many dynamics processing systems offer combinations of the above mentioned possibilities. Such a system might have a level in-level out behavior as depicted in Fig. 22-85.

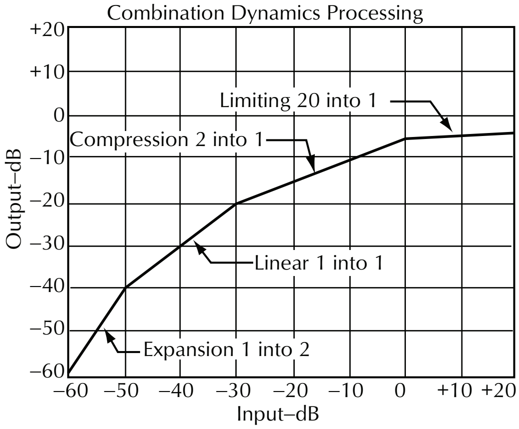


Figure 22-85. Combination dynamics processing.

The processor of Fig. 22-85 features 1:2 expansion for very low level signals, linear behavior for low level signals, compression of 2:1 for intermediate level signals, and limiting of 20:1 as the input approaches high levels. The transition points or thresholds for the different processes as well as the ratios employed in such an instrument might well be under operator control. Additionally, such a device might well feature a linear variable output gain stage that could shift the entire plot either upward or downward. For example an additional gain of 4 dB would place the maximum limited output at 0 dB.

Recording studios might well employ all combinations of dynamics processing, particularly when recording to media of limited dynamic range and or high noise floors. The reproduction of such recordings might well involve complementary dynamics processing. The overall dynamics processing in such instances is referred to as companding. The processing involved in powerful sound reinforcement systems usually consists of linear followed by optional compression and limiting. In venues possessing high ambient noise such as sports arenas or for paging in industrial settings or transportation centers compression is required to insure that the average level of an announcer's voice remains well above the competing noise levels. Limiting is also required in such systems to prevent excessive loudspeaker diaphragm displacement and/or clipping in power amplifiers.

Once the input level exceeds or falls below the threshold for a given processing type, the application or removal of the designated processing is not instantaneous as such behavior would be not only audible but also irritating to the listener. The onset of such processing must be rapid enough to usefully accomplish the goal while at the same time keeping the audible distortion at an acceptable level. Similarly the release of a given processing type must be slow enough not to be noticeable but fast enough to again accomplish the desired objective. The onset and removal behavior is described in terms of attack or release time constants for those analog control systems that employ passive resistive charge or discharge of storage capacitors in the level sensing circuits. More sophisticated systems employ active or constant current charging or discharging techniques. Such systems can be characterized by actual attack and release times. In both instances, the time constants or outright times are usually under operator control. In digital dynamics processing the functions of level sensing as well as the application or removal of a given processing type are accomplished through calculations involving actual sample values and the time histories as presented by sample value sequences. Algorithms are employed that result in operations that mimic or improve upon the behavior of the most successful of analog systems.

As an example of attack behavior of a dynamics processing system consider one in which the operation is linear as long as the input signal has an amplitude less than 1.5 V and goes into hard limiting when the amplitude exceeds 1.5 V. Consider also that this processor is excited by a 1 V amplitude 1 kHz sinusoid from $t = 0$ until $t = 5$ ms at which time the input amplitude becomes 2 V and remains there. The output of such a processor might well appear as depicted in Fig. 22-86.

An examination of Fig. 22-86 reveals that the output of the limiter overshoots the limit value of 1.5 V and then approaches the limit value by means of an exponential decay. The time constant, τ, associated with this decay represents the attack time constant of the limiting action. The envelope of the output is described by

$$V_o = \pm\left(1.5 + 0.5e^{-(t-0.005)/\tau}\right)$$

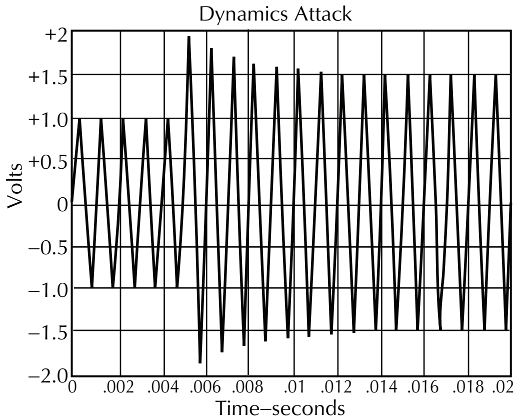


Figure 22-86. Attack behavior of a limiter.

Similarly, consider now that this same processor has been excited for some time prior to a new $t = 0$ by a 2 V amplitude 1 kHz sinusoid and at the new time of $t = 5$ ms the excitation amplitude drops to an amplitude of 1 V and remains there. The recovery behavior from the limiting action of this processor under these conditions appears in Fig. 22-87.

Fig. 22-87 indicates that commencing at $t = 5$ ms the output of the processor undershoots the new required value of 1 V and asymptotically approaches the required value as t grows beyond 5 ms. The envelope of the limiter output is now described by a new equation with τ now being the release time constant.

$$V_o = \pm(1 - 0.5e^{-(t-0.005)/\tau})$$

Many studies have been directed toward the selection of viable values of attack and release times for various types of program material. These studies have involved principally subjective listening tests as the final arbiter is often, "How does it sound?" Many commercial dynamics processors have default values for these parameters based on such studies.

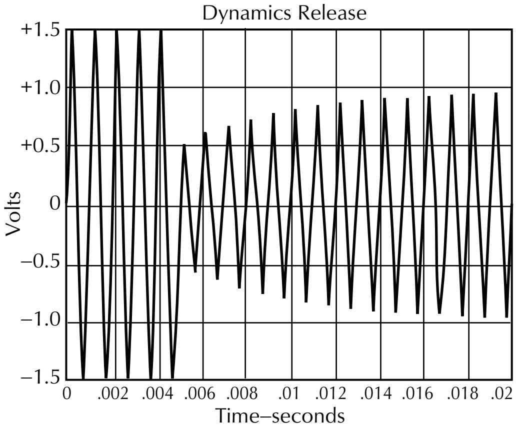


Figure 22-87. Release behavior of a limiter.

Bibliography

G. M. Ballou. *Handbook for Sound Engineers*, 3rd edition. Boston: Focal Press, 2002.

R. J. Higgins. *Digital Signal Processing in VLSI*. Englewoods-Cliffs: Prentice-Hall, 1990.

M. S. Roden. *Introduction to Communication Theory*. New York: Pergamon Press, 1972.

Bernard Sklar. *Digital Communications*. Englewoods Cliffs: Prentice-Hall, 1988.

A. B. Williams and F. J. Taylor. *Electronic Filter Design Handbook*, 2nd edition. New York: McGraw-Hill, 1988.

Chapter 23

Digital Audio Formats and Transports

by Pat Brown

23.1 The Analog Waveform 535
23.2 Quantization 535
  Reconstruction 536
  It’s Not Audio, It’s Data 536
  Data Transport 537
  The Infinite Conveyor 537
  How Fast Must It Be? 538
  Synchronization 539
  Jitter 540
  Latency 540
  Standard Data Formats 540
  The Inner Workings 542
23.3 Digital Signal Processing—DSP 542
23.4 Two Data Camps 543
  Packet-Switched Networks 543
  Network Bandwidth 543
23.5 How Does Ethernet Work? 545
  Who Are You in the Neighborhood? 546
23.6 Ethernet Protocols 546
  The OSI Model 547
  Quality of Service—QoS 549
23.7 An Open Standard 549
23.8 AES3 vs. AoE 549
23.9 Hybrid and Proprietary Systems 549
23.10 Analog vs. Digital Audio 550
  Which Digital Audio “Flavor?” 550
  Learning Digital Audio 551

The intent of this chapter is to present the broad concepts of digital audio sound reinforcement. The inner workings of these systems are amazingly complex, and each process is a field of study within itself. If one invested the time and study to fully understand each part of the digital signal chain, there would be no time left to be an audio practitioner. This bird's eye view is intended to enable one to be conversant in digital audio, and to make informed decisions in selecting and deploying digital audio products. There are a great many excellent resources in-print and on-line for investigating specific parts of the process in more detail.

Much of the information contained in this chapter is the result of a collaboration with Steve Macatee and Brad Benn in producing a digital audio training course. Countless hours spanning several years have gone into wading through the sea of digital theory and identifying the fundamental concepts required to grasp digital audio.

23.1 The Analog Waveform

Air pressure fluctuations can be converted to an analog electrical voltage by use of a pressure-sensitive microphone, Fig. 23-1. This audio waveform can be converted to discrete values for transport or storage. This is the very nature of digital audio, Fig. 23-2.

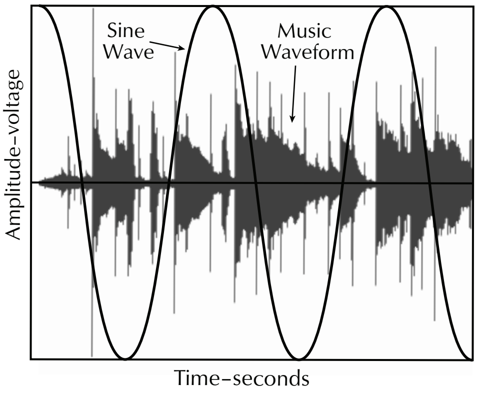


Figure 23-1. Examples of analog waveforms.

Analog waveforms are digitized by voltage sampling at a fixed time interval by an analog-to-digital converter (ADC). The number of available voltage values that may be assigned to each sample is approximately 2^N, where N is the number of bits (ones and zeros) that represent each sample. The more bits, the more available values, and the greater the dynamic range. So, with regard to fidelity "the more bits the better," at least to a practical limit.

According to Nyquist, the sampling rate must be slightly higher than twice the highest frequency present in the waveform to capture all of the audio information that it contains. An anti-aliasing filter may be used to force this condition, as it rejects frequencies for which the selected sample rate is insufficient. This is crucial. For example, if the sampling rate is 48 kHz, the highest audio frequency that can be resolved is approximately one-half, or 24 kHz. An anti-aliasing filter at 24 kHz would assure that no higher frequencies are fed to the ADC.

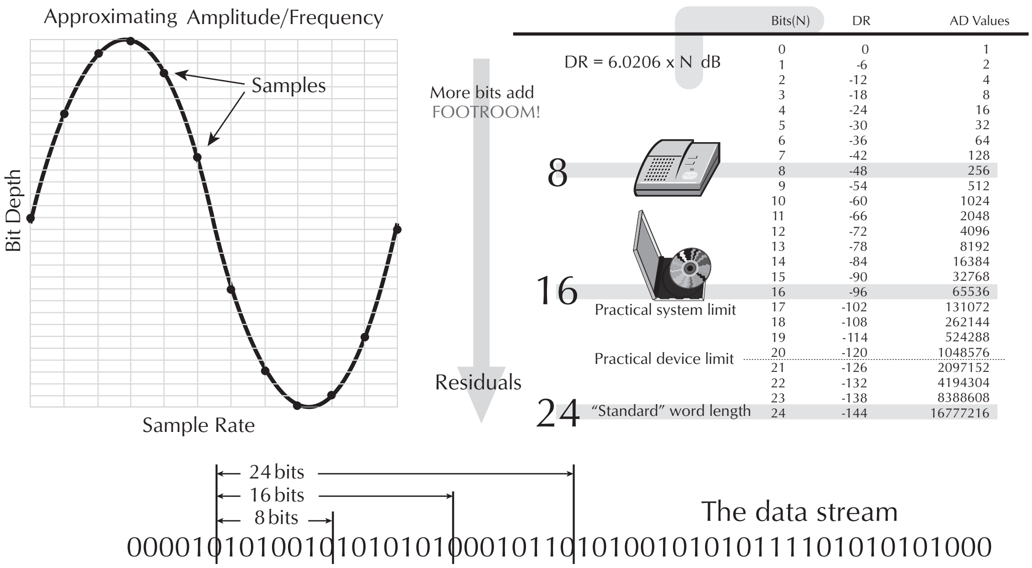
23.2 Quantization

The process of assigning a digital word to each audio sample is known as quantization. A sample of the analog waveform must be rounded to the nearest discrete value as determined by the bit depth. An error is produced from this rounding. The more bits, the lower the quantization error. The accumulated errors form the noise floor of the signal's dynamic range (*DR*).

The two-axis plot in Fig. 23-2 reveals the need for sufficiently small amplitude steps on the Y-axis (bit depth) and a sufficiently short time interval on the X-axis (sample rate) to allow accurate reconstruction of the original analog waveform. "CD quality" of 16 bits and 44.1 kHz sampling rate is well established as the minimum resolution for high fidelity audio reproduction. It is exceeded by the current analog-to-digital converters provided on both consumer and professional audio products. Both sample rate and bit depth can be increased or decreased depending on the desired quality of reproduction. Lower resolution can be acceptable for communication devices, such as telephones and musical greeting cards. Most professional audio devices are CD quality or better.

The string of bits can be divided into 8-bit words. A 16-bit sample is two words. A 24-bit sample is three words. The professional sample rate of 48 kHz is slightly higher than that used for CD quality. Doubling the sample rate adds one musical octave to the audio bandwidth. This is the major reason for 96 kHz and 192 kHz sample rates, the latter producing an audio bandwidth approaching 100 kHz. Increasing the sample rate also reduces the minimum delay step (1 sample) available in digital signal processors, which is an additional motivation for increasing the sample rate beyond what is needed to reproduce the audible spectrum.

A bit depth of 24 bits and a 48 kHz sample rate is commonly presented as 24/48k. This is the default



Bits(N)	DR	AD Values
0	0	1
1	-6	2
2	-12	4
3	-18	8
4	-24	16
5	-30	32
6	-36	64
7	-42	128
8	-48	256
9	-54	512
10	-60	1024
11	-66	2048
12	-72	4096
13	-78	8192
14	-84	16384
15	-90	32768
16	-96	65536
17	-102	131072
18	-108	262144
19	-114	524288
20	-120	1048576
21	-126	2097152
22	-132	4194304
23	-138	8388608
24	-144	16777216

Figure 23-2. The analog waveform must be sampled at a frequency interval. Each sample is assigned a discrete value. The number of available values is determined by the number of bits (word length).

resolution for most professional ADCs. This yields a theoretical dynamic range of approximately 2^{24}, or 144 dB. This is a mathematical resolution. In practice the actual resolution will be far less. A respectable system dynamic range, resulting from stringing together a chain of digital audio devices, is about 100 dB—roughly CD quality.

23.2.1 Reconstruction

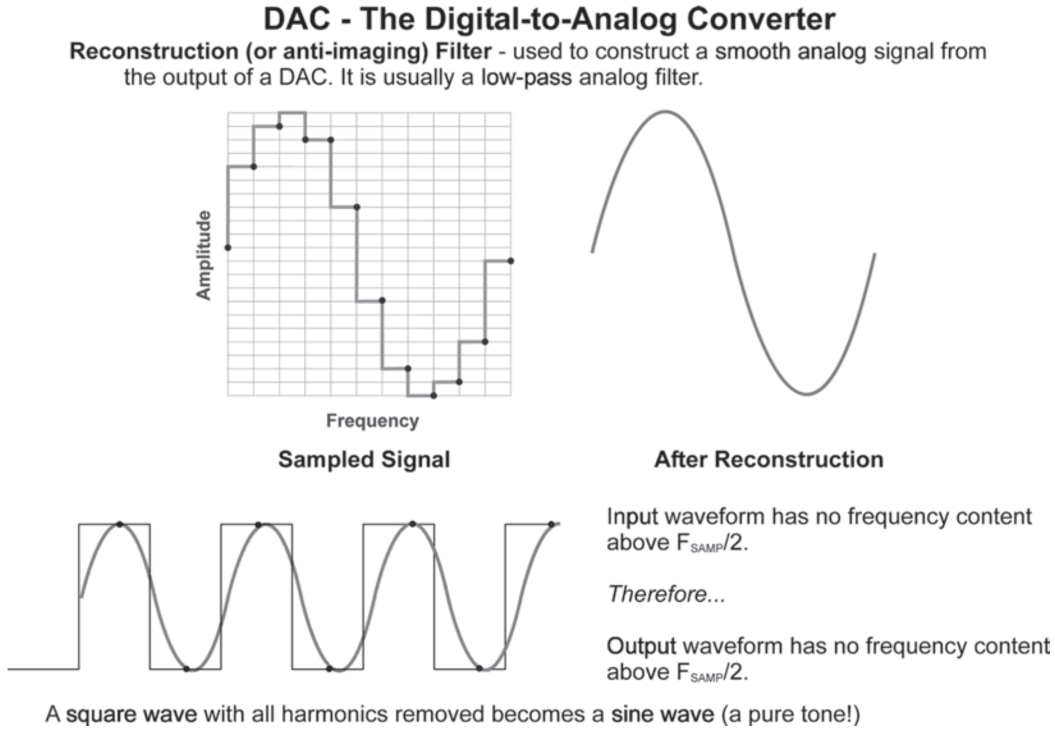
It is non-intuitive that a smooth analog waveform can be recovered from a sampled waveform, especially at high frequencies where there may be only a few samples collected. Digital-to-Analog converters employ a reconstruction filter, usually in the form of a low pass analog filter. The stair steps in the sampled waveform consist of high frequency content outside of the audio band. When it is removed, a smooth analog waveform is the result, Fig. 23-3.

23.2.2 It's Not Audio, It's Data

Once the analog waveform has been converted to a binary code, we leave the continuous, intuitive analog world and enter the deterministic, discrete digital world. While a discontinuity, such as a scratch on a phonograph record, is the worst thing that can befall the analog signal, the digital signal is discontinuous by nature. This is both good and bad. The ability to slice and dice the analog signal into small chunks offers some huge benefits with regard to moving and storing the information. The power of mathematics can be harnessed to process the data in useful ways, including error correction algorithms that can actually replace missing samples. But, unless the data is properly reassembled, the subtle distortions of the analog world give way to devastation in the digital world. The "snow" on an analog TV screen, resulting from low signal strength, is much preferred to the bedlam that results from insufficient digital signal level. This reveals the discrete nature of the digital signal.

A fundamental change in mindset is required to understand what is important and what isn't with regard to digital audio. The interface must be able to distinguish between two states, zero and one. If the signal-to-noise ratio is sufficient for this distinction, there is probably no benefit realized from improving it, if the system passes the signal. While an analog audio system may need 100 dB of dynamic range for pristine reproduction, the digital system need only be able to distinguish between two states to recognize the binary code. In short, high dynamic range analog can be transported over a comparatively low dynamic range medium.

Fig. 23-4 shows the eye pattern test used to check the integrity of the pulses. Note that the AES3 minimum allows significant deterioration of the pulse from the ideal. Unlike analog interfaces, digital interfaces typically either work or they don't. An impaired interface will yield silence, audio with obvious artifacts, or something completely unrecognizable and possibly devastating to a loudspeaker. The "shades of grey" performance improvements in the analog world, possibly realized by the use of



DAC - The Digital-to-Analog Converter

Reconstruction (or anti-imaging) Filter - used to construct a smooth analog signal from the output of a DAC. It is usually a **low-pass** analog filter.

Input waveform has no frequency content above $F_{SAMP}/2$.

Therefore...

Output waveform has no frequency content above $F_{SAMP}/2$.

A **square wave** with all harmonics removed becomes a **sine wave** (a pure tone!)

Figure 23-3. A low pass filter smooths the waveform.

esoteric methods, such boutique cables generally don't apply in the digital world. It's data, not audio.

23.2.3 Data Transport

As with any type of cargo, the transport of digital data involves moving it from point A to point B. This broad term encompasses the many factors that affect data movement. There exist numerous issues and complexities regarding the transport of digital data. As with all engineering practices, compromise is required. The issues at the forefront are:

Data Rate. The rate at which the data flows in bits-per-second.

Channel Count. The number of audio channels interleaved into the digital stream.

Latency. Unintentional and unavoidable delay in transporting the data.

Synchronization. Keeping all audio devices in the system in step.

Formatting. Organization of the data stream so that it can be recognized by the receiver.

As with analog audio theory, the principles of digital audio are not unique to digital audio. I will use a mechanically analogous system to aid in describing the issues that affect data transport.

23.2.4 The Infinite Conveyor

The transport of digital audio data can be visualized as a conveyor belt with one slot or placeholder for each bit. This allows visualization of the serial data stream as it flows through the signal chain. Rather than moving continuously, the conveyor advances one step at a time. The conveyor speed equates to the data rate (or bit rate) of the transport system. The blinding speed of the conveyor (measured in millions of bits-per-second or Mb/s) makes the movement appear to be continuous, even though it moves in discrete steps. The audio words (the payload) must be loaded onto the conveyor at the source and removed at the receiver. The data rate must be high enough to allow additional information beyond the audio words to be interleaved into the data stream. This "metadata" can be used to organize the audio samples into chunks that include information important for recognition, decoding and routing. Metadata may contain information about the data, the data structure, or both. It is data about

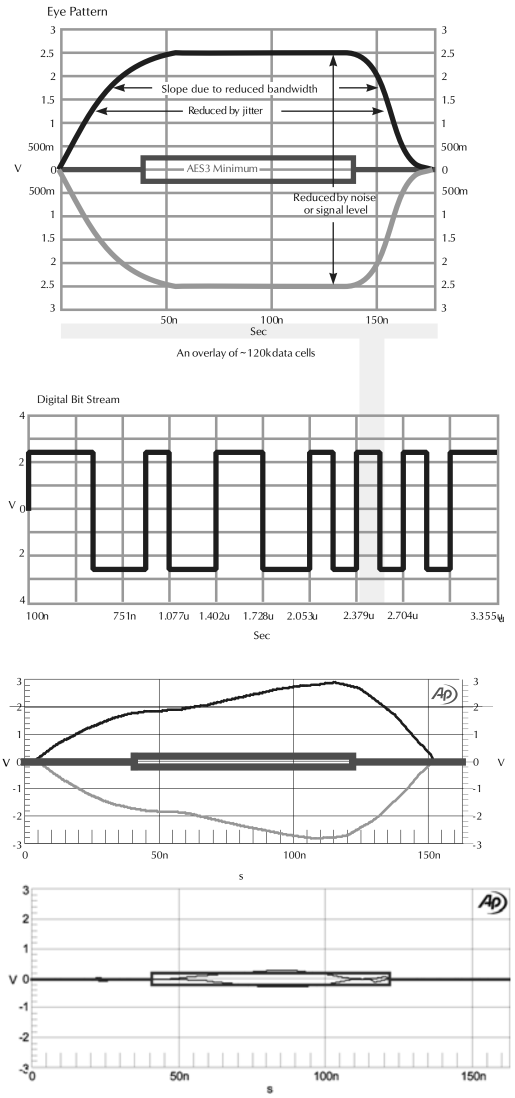


Figure 23-4. The eye pattern test used to check the integrity of the pulses. Shown is the result of the test for 300 ft. of digital audio cable (passed) vs. 300 ft. of foil-shield microphone cable (failed).

data. The process of sampling the data and creating the data stream is called coding.

The data must be decoded at the receiving device to recover the audio information, which may then be reconstructed into an analog waveform which is theoretically identical to the original. The coding and decoding processes are handled by a codec, used here in a general sense. There are many types of codes, the functions and details of which are application-specific. The ADC is complemented by a digital-to-analog converter, or DAC, which is the final step of moving the analog waveform from point A to point B in digital form. The DAC contains a reconstruction filter, which low passes the waveform to remove any residual out-of-band artifacts of the digitization process. This allows a smooth, continuous analog waveform to be recovered from the "stair stepped" analog waveform created from the discrete samples, Fig. 23-5.

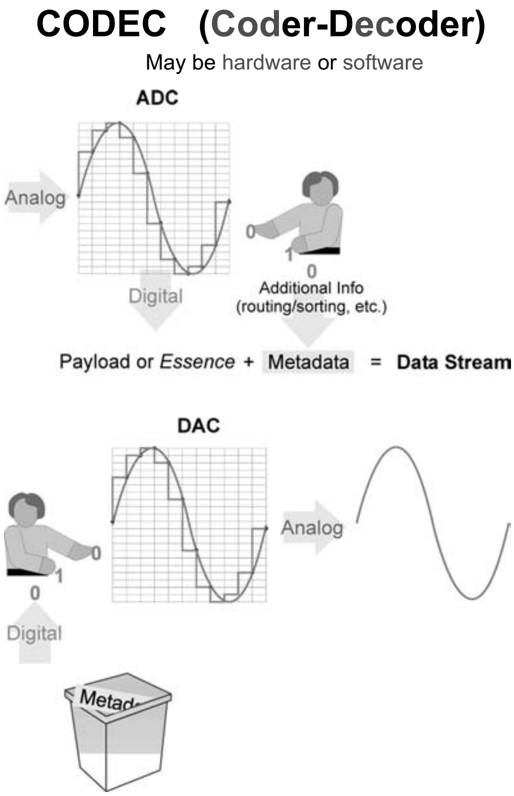

CODEC (Coder-Decoder)
May be hardware or software
ADC
Analog
Digital
0
1
0
Additional Info (routing/sorting, etc.)
Payload or Essence + Metadata = Data Stream
DAC
Analog
Digital
Metadata

Figure 23-5. The codec codes and decodes the analog waveform.

With the distinction made between coding, transport, and decoding, one can see why the increases in bandwidth accomplished over the last several decades have vastly increased the capabilities of digital systems of all types. There exist some amazing possibilities if there is sufficient bandwidth.

23.2.5 How Fast Must It Be?

The minimum required bit rate for the audio samples can be estimated by multiplying the word length in bits times the sample rate, times the number of audio channels. For example, for professional audio quality 16/48k stereo this works out to:

$$16\,(48)(10^3)(2) = 1.54\,\text{Mb/s}$$

For 24/96k data this increases to

$$24\,(96)(10^3)(2) = 4.61\,\text{Mb/s}$$

This simple example in Fig. 23-6 clearly illustrates the trade off involved in increasing the digital resolution. If the data rate is fixed at 100 Mb/s, the number of channels that can be transported is determined by the chosen resolution.

In consumer audio systems, relatively few channels are required. This means that high bandwidth can be used to yield very high resolution digital audio that (arguably) far exceeds the perception abilities of the human auditory system. In professional-audio systems, an unnecessarily high sample rate or bit depth carries the premium of reducing the number of audio channels. Fig. 23-7 shows how human perception relates to sample rate and bit depth.

Note that while digital data is by nature discrete, as the resolution is increased it eventually becomes indistinguishable from analog audio. *Although all physical signals are intrinsically quantized, the error introduced by modeling them as continuous is vanishingly small*. The error is overshadowed by signal noise and instrument inaccuracy. This important statement is obscure in its origins but none the less true.

The data rate must be in excess of what is needed to transport the audio samples alone in order to accommodate the inclusion of metadata, without which the binary code would be unrecognizable to the receiving device. The details requiring standardization include the data stream format (including metadata), the medium (wire type or fiber) and the interface topology, including connectorization.

23.2.6 Synchronization

A major concern in the development of the digital audio interface was how to address the crucial timing that must exist to keep multiple devices and audio channels synchronized. A timing signal, usually at the sampling frequency, is required and is known as word clock. Word clock is the digital equivalent to the timing belt used to synchronize the components in an internal combustion engine, the failure of which can be catastrophic. The timing signal originates at the master clock. This may be a stand-alone device or one of the system components, usually the mixer. The clock signal may be distributed to each digital component (slave) in the signal chain by a dedicated cable. Word clock transport is typically over an unbalanced interface using coaxial cable and connectors. It may be either impedance matched or bridged, since the operating frequency is not extremely high (i.e., 48 kHz). Bridging simplifies word clock distribution in low cost digital audio systems, such as a home recording studio, Fig. 23-8.

Some digital audio signal formats (AES3 and S/PDIF) are self-clocking, meaning that the clock is recovered from the data stream. This allows multiple devices to be synchronized over the same cabling that carries the data stream between them. This simplifies the interconnection of components because only one cable is needed.

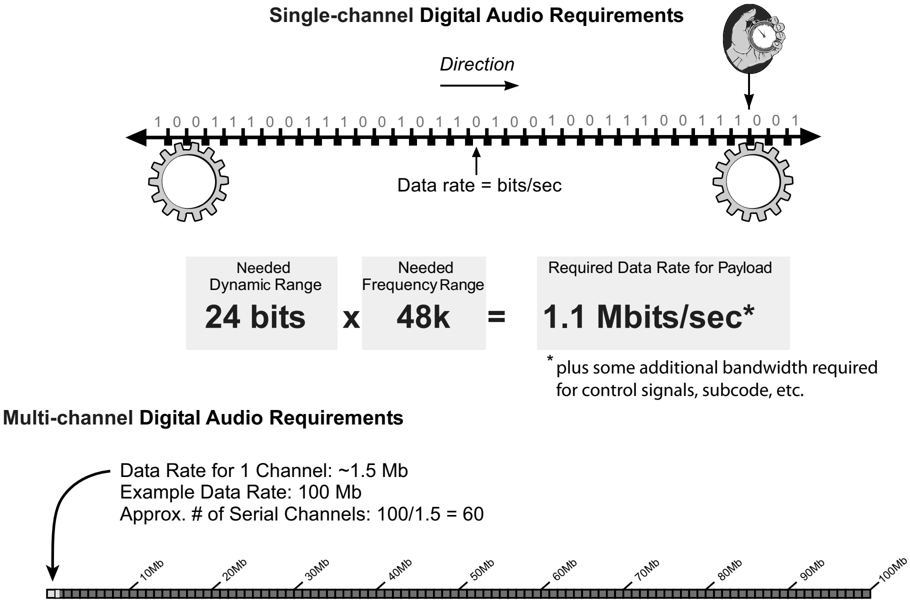

Single-channel Digital Audio Requirements
Direction
1 0 0 1 1 1 0 0 1 1 1 0 0 1 0 1 1 0 1 0 0 1 0 0 1 1 1 0 0 1 1 1 0 0 1
Data rate = bits/sec
Needed Dynamic Range
24 bits
x
Needed Frequency Range
48k
=
Required Data Rate for Payload
1.1 Mbits/sec*
* plus some additional bandwidth required for control signals, subcode, etc.
Multi-channel Digital Audio Requirements
Data Rate for 1 Channel: ~1.5 Mb
Example Data Rate: 100 Mb
Approx. # of Serial Channels: 100/1.5 = 60
10Mb
20Mb
30Mb
40Mb
50Mb
60Mb
70Mb
80Mb
90Mb
100Mb

Figure 23-6. Required data rate for digital audio transport.

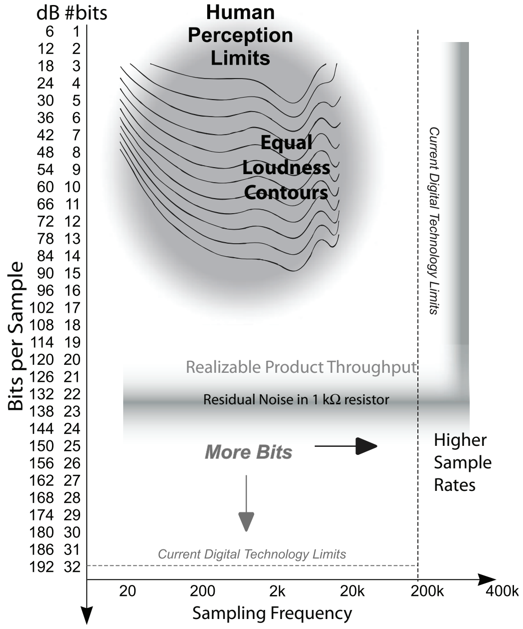


Figure 23-7. Sample rate, bit depth and human perception.

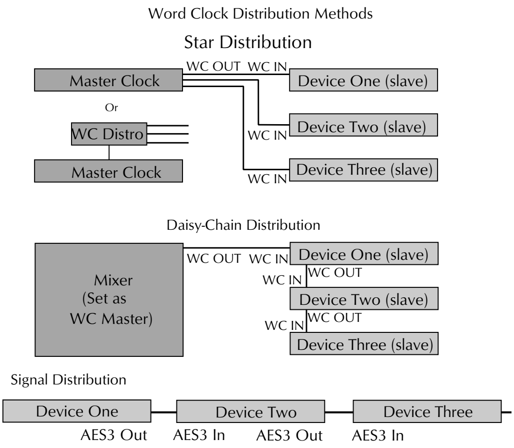

Word Clock Distribution Methods
Star Distribution
Master Clock
WC OUT
WC IN
Device One (slave)
Or
WC Distro
WC IN
Device Two (slave)
Master Clock
WC IN
Device Three (slave)
Daisy-Chain Distribution
Mixer (Set as WC Master)
WC OUT
WC IN
Device One (slave)
WC IN
WC OUT
Device Two (slave)
WC IN
WC OUT
Device Three (slave)
Signal Distribution
Device One
Device Two
Device Three
AES3 Out
AES3 In
AES3 Out
AES3 In

Figure 23-8. Word clock distribution methods.

23.2.7 Jitter

Just as a mechanical conveyor may have some "play" in the gears that skew the timing, the digital audio interface is plagued by jitter, Fig. 23-9. If the jitter becomes high enough, synchronization may be lost. Jitter correction can be included in a digital audio interface. Very high bandwidth transports require very low jitter. It is incumbent to the audio equipment manufacturer to achieve low jitter, and the audio practitioner to preserve it by observing good interfacing practices.

23.2.8 Latency

No matter how fast the bit rate, there is some unintentional and unavoidable delay in the processing and transport of digital audio, referred to as latency. Each device in the signal chain contributes a latency delay. It is cumulative and can eventually become large enough to cause timing problems for a sound system, such as sync between audio and video.

How much is too much? It depends. Even delays of a few milliseconds can cause tonal coloration for a musician playing a saxophone and listening through an in-ear monitor. Audible delay can be perceived when the latency between a live talker and the sound system is 10 ms or so. The latency for a home hi-fi system, can be hundreds of milliseconds with no ill effects. So, the amount of latency that is tolerable depends on whether there is an absolute reference for the listener, such as a talker, singer, stage monitor or video display. Many live sound system designers use a "latency budget" of 10–15 ms.

In contrast, analog systems do not exhibit latency.

23.2.9 Standard Data Formats

There exists many possibilities for organizing the binary data stream. Several standards were devised early in the digital audio development process to avoid having proprietary formats from each manufacturer. A century of analog audio development paved the way for the digital formats. There are some striking similarities between analog and digital formats with regard to implementation and limitations. At the forefront are the need for both consumer and professional versions. The consumer format and interface must be of high fidelity, low cost and simple implementation. Short cable length (a few meters at most) facilitates accomplishment of these objectives. The professional data format and interface adds complexity (and cost) in order to drive longer cables and achieve higher data rates, two goals that tend to be mutually exclusive. The higher data rates can be used to transport large numbers of audio channels, as might be found in venues such as airports and performing arts centers.

The consumer format objectives are currently met by the Sony/Philips Digital Interface, or S/PDIF. The format standardizes the metadata and interface topologies into an essentially "plug and play" interface for transporting high resolution data with low channel counts over short distances. Both wired and optical mediums are available, as are low cost methods for converting between the two. Like its analog counterpart, the wired interface is unbalanced

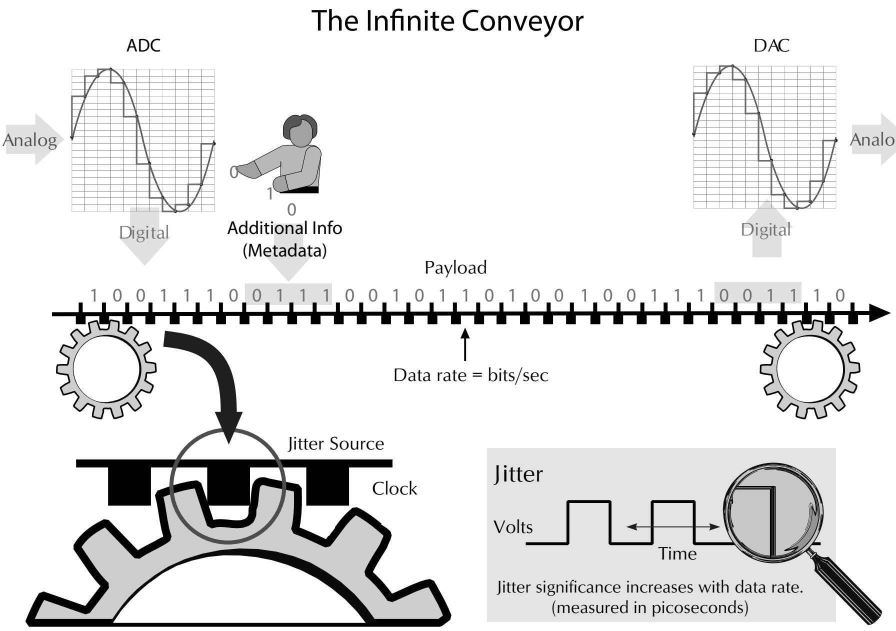

The Infinite Conveyor
ADC
DAC
Analog
Analog
Digital
Digital
Additional Info
(Metadata)
Payload
1 0 0 1 1 1 0 0 1 1 1 0 0 1 0 1 1 0 1 0 0 1 0 0 1 1 1 0 0 1 1 1 0
Data rate = bits/sec
Jitter Source
Clock
Jitter
Volts
Time
Jitter significance increases with data rate.
(measured in picoseconds)

Figure 23-9. Jitter can upset the delicate timing required for data transport.

and utilizes coaxial cable and 2-conductor connectors. The interface is impedance-matched at 75 Ω and the signal voltage is about 1 $V_{p\text{-}p}$.

The professional format objectives are met by AES3 and its nearly identical European sibling, AES-EBU. The data format is nearly identical to S/PDIF, the difference being in some details regarding the metadata. Like its professional analog counterpart, the wired interface is balanced electronically and requires a twisted-pair cable for transport. Connectors are 3-conductor to accommodate an optional cable shield. The interface is impedance-matched at 110 Ω and the signal voltage is 3–7 $V_{p\text{-}p}$. Conversion between AES3 and S/PDIF can be accomplished with a passive network.

Other professional and consumer formats exist, some of which support higher channel counts. While the details may be different, in concept they are similar to AES3 and S/PDIF, Fig. 23-10.

Both S/PDIF and AES3 can be carried over cables designed for analog interfaces, which was one of the requisites that influenced the standardization of each interface. Since the interface must only recognize two states—zero and one—the wiring characteristics are actually more forgiving than for analog signals until the data rate becomes very high or the cable very long. "Digital audio cable" is designed for lower capacitance per unit length than wire designed primarily for analog use. This can become a factor as the 100 meter limit for the basic AES3 interface are approached or exceeded. Yes, a "mic cable" can carry digital audio without degradation, ultimately limited by cable length. The high twist ratio of Category cable (i.e., CAT 5 or CAT 6) makes it suitable for carrying AES3 data. As with balanced analog interfaces, the cable shield is not

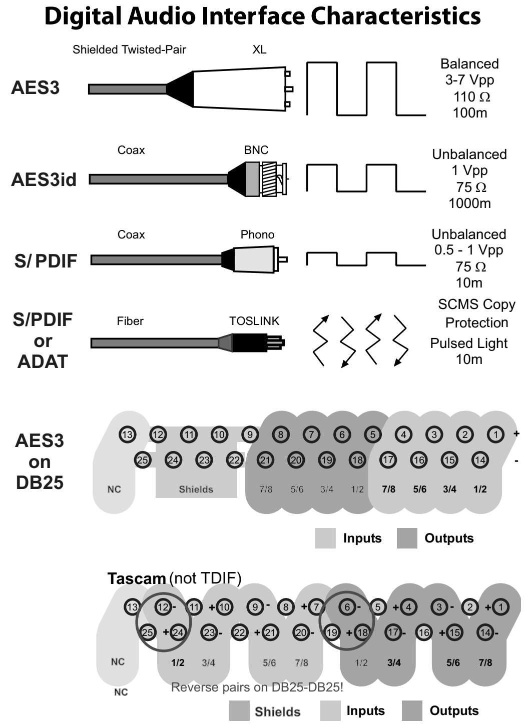

Digital Audio Interface Characteristics
Shielded Twisted-Pair
XL
AES3
Balanced
3-7 Vpp
110 Ω
100m
Coax
BNC
AES3id
Unbalanced
1 Vpp
75 Ω
1000m
Coax
Phono
S/ PDIF
Unbalanced
0.5 - 1 Vpp
75 Ω
10m
S/PDIF
or
ADAT
Fiber
TOSLINK
SCMS Copy
Protection
Pulsed Light
10m
AES3
on
DB25
NC
Shields
7/8 5/6 3/4 1/2 7/8 5/6 3/4 1/2
Inputs
Outputs
Tascam (not TDIF)
NC
1/2 3/4 5/6 7/8 1/2 3/4 5/6 7/8
NC
Reverse pairs on DB25-DB25!
Shields
Inputs
Outputs

Figure 23-10. Digital audio interface characteristics.

required for signal flow for AES3, but serves the purpose of:

1. Containing the digital audio signal field.
2. Rejecting extraneous fields.
3. Reducing the chassis voltage difference between source and load devices.

The electrical properties of video signal distribution were found to work well for AES3 audio. These include an unbalanced 75 Ω interface, coaxial cable and active or passive splitters. The AES3id Standard provides the details for transport over coaxial cable.

23.2.10 The Inner Workings

Both AES3 and S/PDIF are highly complex, consisting of two interleaved audio channels organized into blocks, frames and subframes, Fig. 23-11. There are always two channels present, even if both contain the same information (mono).This convention was no doubt influenced by the stereo playback system popular in consumer playback systems and recording studios.

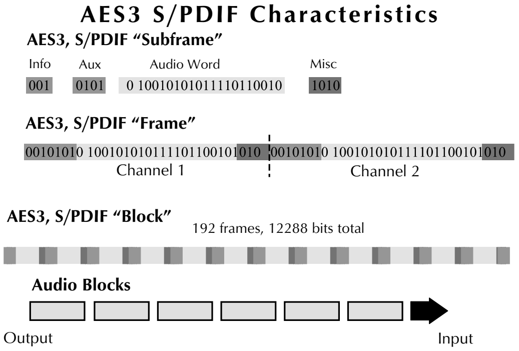

AES3 S/PDIF Characteristics
AES3, S/PDIF "Subframe"
Info
Aux
Audio Word
Misc
001
0101
0 10010101011110110010
1010
AES3, S/PDIF "Frame"
00101010 100101010111101100101010
00101010 100101010111101100101010
Channel 1
Channel 2
AES3, S/PDIF "Block"
192 frames, 12288 bits total
Audio Blocks
Output
Input

Figure 23-11. Details of the AES3 and S/PDIF formats.

Returning to the conveyor belt analogy, the data stream is flowing, even with no audio present. A clock signal is recovered from the stream and is sensed by the interface and used to indicate a successful connection, known as achieving "lock." When lock is achieved, data is flowing. Some audio devices provide visual indication that lock has been achieved, as well as metering to observe the audio being transported.

Sophisticated instrumentation is required for detailed observation and analysis of the data stream. In most cases, the audio technician needs only to know if the clock signal and audio data are present, which greatly simplifies the required instrumentation. Some products actually test the cable when a connection is made, and visually indicate any faults. A hand-held instrument can detect the clock (lock), determine the sample rate and strip off the audio for listening or measurement. It is the digital technician's equivalent to the telephone butt set.

AES formats have expanded to accommodate microphones (AES42) and high channel counts (AES50).

23.3 Digital Signal Processing—DSP

One of the greatest strengths of digital audio is with regard to signal processing. DSP has revolutionized the sound reinforcement industry. It would be difficult today to find a sound system that does not include a DSP. Once the analog waveform is represented numerically, the power of mathematics can be used to process it in many ways.

Some digital signal processes merely emulate their analog counterparts. The effect of the DSP on the signal is indistinguishable from the analog process, except for latency. This type of digital filter is called an Infinite Impulse Response, or IIR filter. Fig. 23-12 shows how an analog low pass filter could be implemented digitally using mathematics. Note that both filters are recursive, meaning that some of the output is fed back to the input, making future sample values depend on past sample values. In theory, if fed an impulse such a filter never decays to zero, hence the name. Most crossover and filter blocks in a DSP use IIR filters. An IIR filter emulates an analog filter in both magnitude and phase response. This is important for loudspeaker equalization applications where it is desirable to correct minimum phase anomalies in the loudspeaker's response. In this application, the phase shift produced by an analog filter, or an IIR digital filter is desirable, since it is the conjugate of the phase shift in the response bump or dip being corrected.

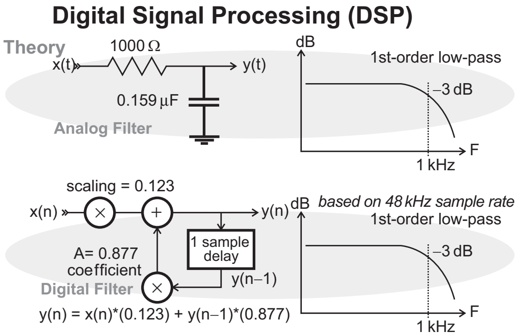

Digital Signal Processing (DSP)
Theory
1000 Ω
x(t)
y(t)
0.159 µF
Analog Filter
dB
1st-order low-pass
–3 dB
F
1 kHz
scaling = 0.123
x(n)
y(n)
1 sample delay
A= 0.877 coefficient
y(n–1)
Digital Filter
y(n) = x(n)*(0.123) + y(n–1)*(0.877)
dB
based on 48 kHz sample rate
1st-order low-pass
–3 dB
F
1 kHz

Figure 23-12. Implementation of an IIR digital filter. (*Courtesy Steve Macatee.*)

A second type of digital filter is the FIR, or Finite Impulse Response filter. An FIR filter is a fixed length impulse response that is convolved with the digital audio waveform, imparting its characteristics upon it. Convolution can be used to encode anechoic program material with the measured or simulated impulse response of a room, allowing evaluation of the room's acoustical characteristics. It could also be used to convolve a high pass response onto the digital audio data, such as might be used in a crossover network. FIR filters have the interesting characteristic of being non-minimum phase, meaning that there is no predictable relationship between the magnitude and phase response of the filter. This can allow the magnitude response to be modified without changing the phase response. A practical application is the formation of a very steep high pass filter for protecting a high frequency driver. A steep slope produced with an analog or IIR filter would exhibit significant phase shift which could be perceived as time smear by a listener. An FIR filter could create the steep slope without causing phase shift. The "Linear Phase Brickwall" filter has become a favorite of loudspeaker designers. But, there is no free lunch, as FIRs tend to have longer latency than IIR filters, Fig. 23-13.

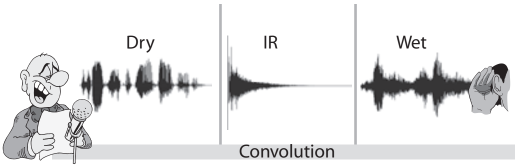


Figure 23-13. Convolving an impulse response of finite length (.wav file) with dry program material.

23.4 Two Data Camps

Of course the need to transport digital information efficiently between sources and receivers extends well beyond audio. Bits are bits, so there is the potential to utilize other existing data delivery schemes for the transport of audio. The massive R & D and investment that has produced the wide bandwidth local area computer network (LAN) for offices can serve as a digital audio transport system.

There are two dominant means for the transport of digital audio in professional systems. The first is the use of audio industry-specific formats and interfaces. These include AES3 and its variants. The second is the use of Ethernet technology, as borrowed from the computer networking industry. Several versions of Audio-over-Ethernet (AoE) provide viable alternatives to the dedicated audio formats.

23.4.1 Packet-Switched Networks

"Packet-switched network" is a general term for a digital communications network that groups all transmitted data, irrespective of content, type, or structure into suitably sized blocks, called packets, or datagrams. It is analogous to a shipping company that uses the same carton or container to achieve standardization, without regard for the content. While in some ways similar to the frames used to transport AES3 data, a packet can contain non-audio data, allowing audio packets to be interspersed with other, non-audio packets.

Conceptually, if we have an extremely fast, accurate conveyor, and a means of accurately on-loading and off-loading the ones and zeros, analog information of all types can be transported to one or more destinations. It is not unlike a train made up of a multitude of cars carrying various types of cargo from one destination to another. There are sources and destinations, as well as terminals and rail yards along the way that connect them. There is a label on each package to indicate its origin and destination. The key is to find a route from point A to point B, Fig. 23-14.

Ethernet has emerged to become the dominant type of packet-switched network used for the local area network (LAN). An Ethernet network is comprised of nodes interconnected through switches. The nodes may be computers, printers or other office appliances. A switch (or bridge) is a highly intelligent multi-port device that routes the data between the nodes. Nodes and switches are connected in a star topology. Each node has a unique identifier to distinguish it from other nodes on the network. This may be a physical Media Access Control (MAC) address, which all networkable devices are given at the time of manufacture. If the device resides on the Internet, there will be an additional Internet Protocol (IP) address, that can be user-configured for a specific system or application. An Ethernet switch compiles a list of nodes connected to its ports. This allows it to route incoming packets to their intended destination, Fig. 23-15.

23.4.2 Network Bandwidth

We are now transitioning from the discussion of dedicated digital audio formats to the discussion of digital data transport in general. We have covered many of the principles already, and will now expand the discussion from point-to-point interfacing to networks. The bandwidth of a network describes the data rate or bit rate. A 10Base-T Ethernet network runs at 10 Mb/s and uses twisted-pair (CAT 5 or

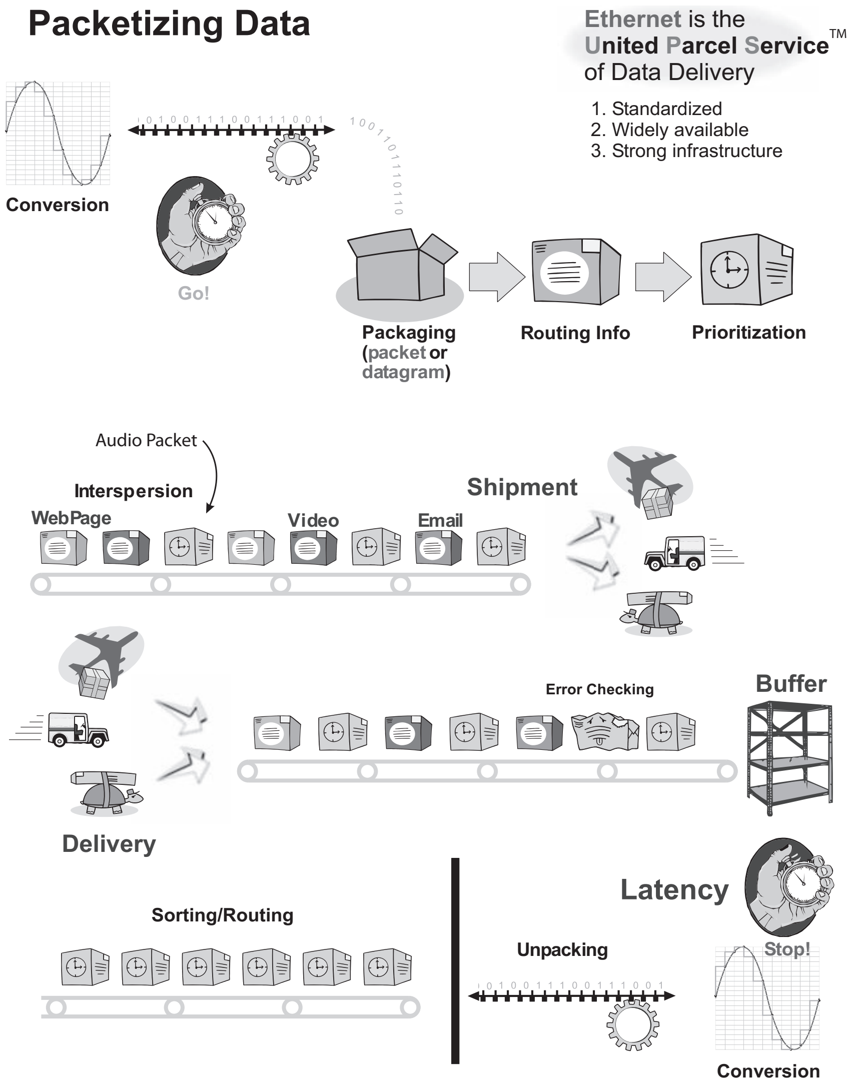

Packetizing Data
Ethernet is the
United Parcel Service™
of Data Delivery
1. Standardized
2. Widely available
3. Strong infrastructure
Conversion
0 1 0 0 1 1 1 0 0 1 1 1 0 0 1
1001101110110
Go!
Packaging
(packet or
datagram)
Routing Info
Prioritization
Audio Packet
Interspersion
WebPage
Video
Email
Shipment
Error Checking
Buffer
Delivery
Latency
Sorting/Routing
Unpacking
Stop!
0 1 0 0 1 1 1 0 0 1 1 1 0 0 1
Conversion

Figure 23-14. Using Ethernet as a means of audio transport.

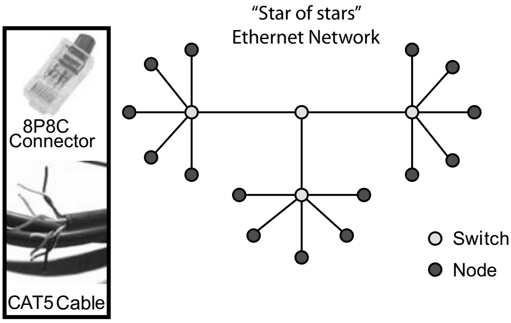

8P8C
Connector
CAT5 Cable
"Star of stars"
Ethernet Network
Switch
Node

Figure 23-15. An Ethernet network. Any node can communicate with the other nodes. The network traffic is routed by a switch (bridge).

higher) cabling. Since a full-bandwidth audio channel requires a data rate of about 1.5 Mb/s, several channels could be transported over a 10Base-T network. Technological advancements produce increased bandwidth, with typical bandwidth leaps of an order-of-magnitude (ten-fold) occurring every few years. With 100Base-T Ethernet (known as Fast Ethernet), AoE becomes quite interesting to system designers, as the channel count for full bandwidth audio approaches 60. 1000Base-T (Gigabit) networks promise to take this to nearly 600, making the construction of very large, complex audio systems with sophisticated routing a possibility, Fig. 23-16.

High bandwidth is a two-edged sword. As the data rate increases, so do the details regarding the cabling and connectors. Category cable (i.e. CAT 5,CAT 5e, CAT 6) is a twisted-pair cable rated based on crosstalk and system noise. 10Base-T and 100Base-T networks can be somewhat forgiving

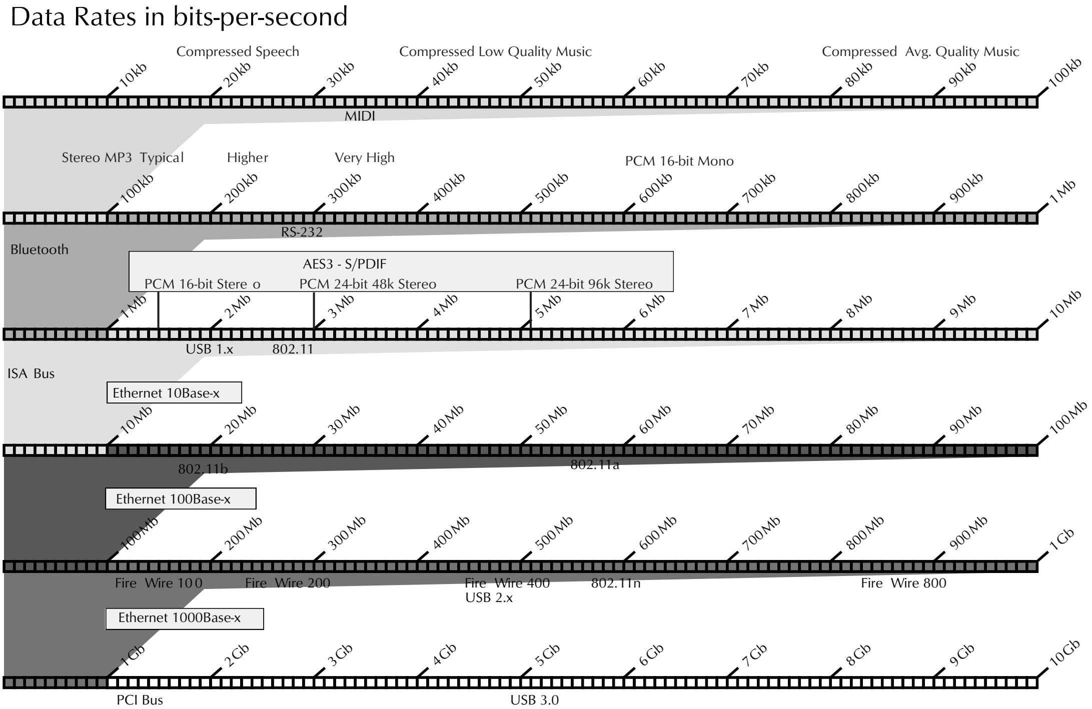


Figure 23-16. Data rates and digital audio formats.

regarding cable installation and routing details. Their needs are satisfied by Category 5 (CAT 5) cables. A CAT 5 cable consists of four 24 AWG twisted pairs sharing a common jacket. It is terminated by the 8P8C connector, often referred to by a telephone company wiring topology for which it is used–RJ45. The "CAT 5 cable" has become ubiquitous, sometimes being used to carry proprietary digital data formats and even analog audio.

As the data rate (bandwidth) increases, the cable requirements become more stringent. CAT 6 cabling utilizes a larger wire gauge (22 AWG) to increase its usable bandwidth. The impedance change caused by a sharp bend or a tight cable tie can produce reflections and reduce the data integrity, requiring that packets be resent by error correction algorithms if the protocol supports it, thereby reducing the useful bandwidth. In severe cases, drop-outs may occur. 1000Base-T networks require CAT 6 cabling.

The highest potential bandwidth is afforded by fiber optic cabling, which uses pulsed light to convey the bits rather than electrical potential. The light can travel a long distance before it becomes dispersed within the medium. Multi-mode fiber has a large diameter relative to the signal wavelength. The light reflects geometrically down the fiber and is recovered at the receiving device. Single-mode fiber has a small diameter relative to the signal wavelength. It is analogous to the acoustical plane wave tube. Single-mode fiber delivers the highest possible bandwidth (or longer distance) at an increased cost over multi-mode fiber.

Because the effect of dispersion increases with the length of the fiber, a fiber transmission system is often characterized by its bandwidth–distance product, usually expressed in units of MHz×km. This value is a product of bandwidth and distance because there is a trade off between the bandwidth of the signal and the distance it can be carried. For example, a common multi-mode fiber with bandwidth–distance product of 500 MHz×km could carry a 500 MHz signal for 1 km or a 1000 MHz signal for 0.5 km.*

Fiber optic terminations can be especially problematic. While analog cables can be checked by a continuity test, high speed data networks require specialized (and expensive) instrumentation for testing. A third party may be contracted to certify the cabling due to the expense of the instrumentation and the required expertise.

23.5 How Does Ethernet Work?

We have now transitioned away from audio, in a textbook devoted to audio. There exist a vast number of resources regarding data transport over Ethernet networks. It is a field of study and expertise in and of itself, in which the experts may need to understand little or nothing about the actual data flowing around the network. Their job is to ensure that data that is streamed onto the network is delivered to its destination intact, and to minimize the down time of the network. The vast majority of audio practitioners only need a birds-eye view of how Ethernet networks work, which I will attempt to establish in the following sections.

---

* Wikipedia—Fiber Optic Communication.

23.5.1 Who Are You in the Neighborhood?

The MAC address is unique for every Ethernet device in existence (node). It is assigned at the time of manufacture, and cannot be changed. This is called the physical address of the node. Depending on the network type, devices may be given an additional virtual address, called the Internet Protocol (IP) address.

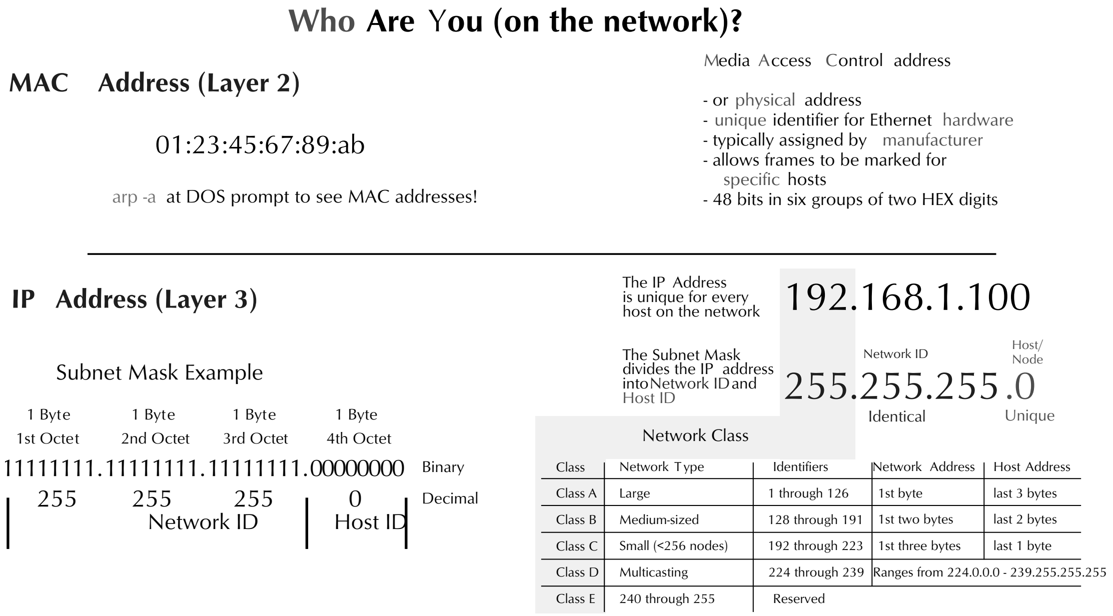
An Ethernet network can be envisioned as a neighborhood comprised of X streets with Y houses on each street. Each street needs a unique identifier (the Network ID), as does each house (the Host ID). Only then can datagrams be properly routed from house to house (node to node). The IP address is a 32-bit string that is divided into four octets, Fig. 23-17. Part of this string designates the street name (or subnet). This is the Network ID. The remaining part designates the house number. This is the Host ID. A subnet mask determines which bits are the street name and which bits are the house number. The subnet mask is a 32-bit string. The left most bits will be "1" and the right most bits will be "0." It tells the network how to interpret the IP address of a node. For example, if the first 24 bits are ones and the final 8 bits are zeroes, this tells the network to interpret the IP address as the first three octets signifying the network ID and the final octet signifying the Host ID. The subnet mask is usually represented by decimal rather than binary values for simplicity.

The Network ID must be the same for every node on the network. The Host ID must be unique for every node on the network.

For example, how I subdivide the 32-bit IP address could allow a network with lots of streets with a few houses on each, or just a few streets with many houses on each. The IP address can be assigned manually. While this can work for small networks, a more sophisticated approach is needed for large networks. A Dynamic Host Configuration Protocol (DHCP) server can reside on a network and dole out unique IP addresses for each device that is connected.

23.6 Ethernet Protocols

In computing, a protocol is a set of rules governing the exchange or transmission of data between devices. Many possibilities exist. An Audio-over-Ethernet (AoE) transport system breaks the digital audio stream up and transports it in packets.


Who Are You (on the network)?
MAC Address (Layer 2)
01:23:45:67:89:ab
arp -a at DOS prompt to see MAC addresses!
Media Access Control address
- or physical address
- unique identifier for Ethernet hardware
- typically assigned by manufacturer
- allows frames to be marked for specific hosts
- 48 bits in six groups of two HEX digits
IP Address (Layer 3)
Subnet Mask Example
1 Byte
1st Octet
1 Byte
2nd Octet
1 Byte
3rd Octet
1 Byte
4th Octet
11111111.11111111.11111111.00000000
Binary
255 255 255
Network ID
0
Host ID
Decimal
The IP Address is unique for every host on the network
192.168.1.100
The Subnet Mask divides the IP address into Network ID and Host ID
Network ID
Host/ Node
255.255.255.0
Identical
Unique
Network Class

Class	Network Type	Identifiers	Network Address	Host Address
Class A	Large	1 through 126	1st byte	last 3 bytes
Class B	Medium-sized	128 through 191	1st two bytes	last 2 bytes
Class C	Small (<256 nodes)	192 through 223	1st three bytes	last 1 byte
Class D	Multicasting	224 through 239	Ranges from 224.0.0.0 - 239.255.255.255	
Class E	240 through 255	Reserved		

Figure 23-17. Network node identifiers.

The data rates are high enough to allow non-audio packets such as DSP control or even video to be interleaved with the audio packets. AoE protocols follow the Open Systems Interconnection (OSI) layer model for data transport. The OSI model provides a worldwide, abstract, theoretical framework for understanding and discussing data transport. A general understanding of the OSI reference model provides the framework for understanding the differences between the AoE protocols, as well as problem solving.

23.6.1 The OSI Model

The workings of a computer network can be broken down into layers, with each layer describing some aspect of how the network works. Fig. 23-18 shows the Open Systems Interconnection (OSI) 5-layer model, which simplifies the original 7-layer model by combining some of the upper layers. I'll provide an overview of the layers used for audio transport. The layers are usually presented and visualized from the bottom up, the variable between protocols being how many layers (if any) above Layer 1 (the physical cabling) are used.

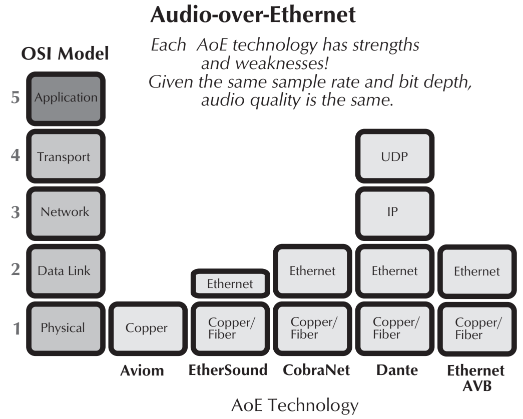


Figure 23-18. The OSI Reference Model, and the layers used by AoE protocols.

All of the data on an Ethernet network is in the form of packets. A "packet sniffer" is a software application that allows the packet activity to be viewed and analyzed. All packets that flow over the network contain routing information. The layers can be thought of as the address label on a physical shipment. The higher the layer, the more information included on the label.

Layer 5 - Application.
Layer 4 - Transport.
Layer 3 - Network.
Layer 2 - Data Link.
Layer 1 - Physical.

Layer 1, the physical layer, describes the medium over which the data travels. This could be copper wire, fiber or even radio frequency transmission. AoE that uses only Layer 1 is not technically Ethernet. It just uses the media designed for Ethernet. The user is shielded from the complexities of the inner workings of the network, and is able to realize many of the strengths of AoE in an essentially "plug and play" implementation. Aviom™ is a commercial example.

Layer 2 protocols transmit Ethernet packets containing audio samples from node to node based on each device's Media Access Control (MAC) address. This is a unique identifier assigned to the audio product by the manufacturer, and is sometimes called its physical address. It is like the street address of a house, except that if the house is moved the address goes with it. Cobranet™ and Ethersound™ are commercial examples.

Layer 3 protocols can route audio data to a device's Internet Protocol (IP) address, a virtual address that can be assigned by the user. The IP address must be unique on a given network, so audio devices that have the same IP address must reside on different networks. If connected to the same network, a conflict occurs that must be resolved for data to flow. Routing audio data via the IP address adds another level of sophistication and complexity to the transport.

Audinate's Dante™ is a commercial example. Ethernet AVB is an open standard example.

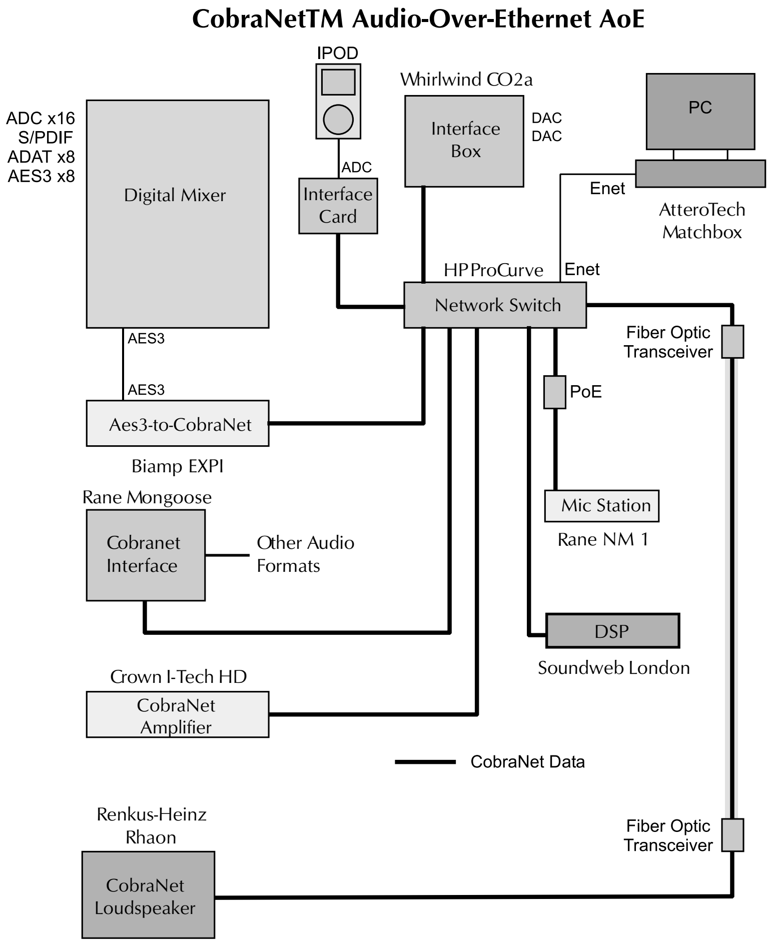
These Standards are either licensed or proprietary, where networking technology is a product unto itself and forms a revenue stream for the company who developed it. They cannot be intermixed nor can a device using one AoE format directly communicate with one using another without a protocol converter. It is very important to research the strengths and weakness of each to decide which is most appropriate for a given application.

An Example

Let's use the OSI model to discuss the creation of a digital audio network. The first decision is Layer 1. What media do we wish to use to interconnect the components? Our choices may include fiber, CAT-5 or coaxial cable. It is decided that CAT-5 will be used, due to its widespread availability and low cost. It is decided that Cobranet™, a Layer 2 protocol will be used for transporting the digital data. We select some products that support Cobranet and connect them to the network, and follow the manufacturer's instructions for routing the data. The MAC address of each device identifies it on the network. Some of the audio devices require control from a PC. The control data will be on Layer 3, which means that each device will be given a unique IP address, either manually or by a DHCP server. The controlling PC is given IP address 192.168.0.100 with a subnet mask of 255.255.255.0. This means that the devices that it controls will be given an IP address with the same subnet mask (255.255.255.0), the same network ID (192.168.0.nnn) and a unique Host ID (.nnn). This enables them to communicate on the network. This will allow 254 devices to reside on the network, which is the number of unique Host IDs allowed by this subnet mask.

It's now time to assess the bandwidth requirements of our audio system to determine if it can reside on the customer's existing network. It is determined that it can, so we request that a Virtual Local Area Network (VLAN) be established by the customer's IT department to keep our audio traffic off of the main network, and to keep the main network traffic out of our audio. Their network switches reside on Layer 3,which is required for the establishment of a VLAN, Fig. 23-19.

The OSI Reference Model has allowed us to discuss the establishment of the digital audio network. We had several options for each layer, and made our choices based on the specific application. A future system design may use fiber on Layer 1, and route both the audio and control data on Layer 3, using an IP address to identify each device in lieu of the MAC address. A Layer 3 protocol


CobraNetTM Audio-Over-Ethernet AoE
IPOD
Whirlwind CO2a
PC
ADC x16
S/PDIF
ADAT x8
AES3 x8
Digital Mixer
ADC
Interface Card
Interface Box
DAC
DAC
Enet
AtteroTech Matchbox
HP ProCurve
Enet
Network Switch
AES3
AES3
Fiber Optic Transceiver
PoE
Aes3-to-CobraNet
Biamp EXPI
Rane Mongoose
Mic Station
Rane NM 1
Cobranet Interface
Other Audio Formats
DSP
Soundweb London
Crown I-Tech HD
CobraNet Amplifier
CobraNet Data
Renkus-Heinz Rhaon
Fiber Optic Transceiver
CobraNet Loudspeaker

Anything can connect anywhere - connections made by addressing!

Figure 23-19. A digital audio network. The nodes that use CAT-5 can be separate by up to 100m. The fiber links allow much greater distances.

must be used, such as Audinate's Dante™. We could then create separate subnets for each building on a campus, allowing many audio networks to function independently over the same Layer 1 media. A router could be used to allow the subnets to pass data between them. Many possibilities exist.

23.6.2 Quality of Service—QoS

A major challenge regarding the use of Ethernet as an audio transport is to ensure that the continuous audio waveform can be reconstructed without discontinuities, and with minimal latency. The audio packets are flowing over the network interspersed with non-audio packets. Small delays that occur when downloading files or receiving emails are generally acceptable. We give no thought to waiting a few extra seconds for the information if the network is busy. This is completely unacceptable for real-time audio streaming, where delayed packets could produce a discontinuity or drop-out in the signal. The umbrella term for the discussion of data packet priority and other delivery factors is Quality of Service (QoS).

AoE can be given a high QoS through the use of a dedicated network for the audio system. Ethernet switches (bridges) are relatively inexpensive these days, but unfortunately the installation of separate cabling for the audio network may not be. It is extremely attractive to the venue owner to utilize their existing LAN for audio, which means that it may also be carrying data for cash registers, credit card transactions, emails, web surfing, etc. A Virtual Local Area Network (VLAN) can segment part of an existing LAN, allowing it to function as a stand-alone network. It is common for audio system designers to request (from the IT staff of a facility) the establishment of a VLAN for AoE. The sound system designer will be queried regarding the bandwidth requirements of their design by the IT department of the facility.

23.7 An Open Standard

An open standard is one that is developed and agreed upon by a standards organization, such as the IEEE or AES. Ethernet AVB is an IEEE Standard that bridges audio and video into the same data stream. It is a Layer 3 technology that addresses some important issues for audio and video transport, such as the need for high QoS, to give these packets higher priority for delivery over the network. In effect, Ethernet AVB establishes its own "protected cloud," giving high QoS to audio and video data. Ethernet AVB is still in development but is expected to emerge as a dominant AoE protocol.

It is highly non-intuitive for audio to be sliced up at one end and pieced back together at the other end to reconstruct a continuous analog waveform, yet this is precisely how AoE works. There exists no intuitive connection between transporting audio data over Ethernet and the use of point-to-point transport schemes like AES3 or analog. The need to acquire IT skills is a major downside of AoE for the audio technician, who is already tasked with mastering many disciplines. This preserves the attractiveness of audio industry-specific formats, such as AES3 for many applications.

Given the same sample rate and bit depth, all AoE technologies sound the same. As such, there is no reason to migrate from one AoE protocol to another to achieve better sound quality. There are differences in latency that could be a factor for some applications. The designer of a digital audio sound system must know their "latency budget" and stay within it to avoid timing problems.

23.8 AES3 vs. AoE

Which will win? Hopefully neither. There's a strong upside for having industry-specific protocols for digital audio transport. Advantages include simpler implementation, analog-like cabling and connectors, and a similar point to-point mentality as analog interfaces. AoE, on the other hand, rides on the huge investment made by non-audio industries. Audio can ride on the existing network backbone in large facilities, sectioned off as a dedicated VLAN. Low-cost switches and cabling allow the economical construction of dedicated audio networks, where audio can be routed to any address on the network. While implementation may require significant networking knowledge and skills, this requisite can be expected to diminish with time. Perhaps the greatest benefit offered by AoE is its flexible routing capabilities. One need no longer to think of signal flow being confined to serially connected devices. A signal processor can be connected anywhere on the network to process a loudspeaker located anywhere else on the network.

23.9 Hybrid and Proprietary Systems

Many engineering and medical solutions use the strengths of multiple approaches to achieve a result unattainable with any single method. This is also true of digital audio systems. A modern audio product may provide analog inputs to accommodate the sundry program sources that must interface to the system. Add-on modules may provide additional analog or digital audio inputs that connect back to the main unit using a digital interface. This allows a system to be highly customized, yet retain the simplicity of point-to-point connections using network cabling. Many of the strengths of digital audio can be realized while shielding the user from the technical details, Fig. 23-20.

23.10 Analog vs. Digital Audio

If analog audio is from Mars, digital audio is from Venus. Many variables, some subtle and some not, affect the analog waveform. This has produced a touchy, feely "shades of grey" medium in which artistic preference may have as much influence as "technical correctness." Analog audio can "sort of" work, and systems that are far from optimal may still fulfill their intended purpose. Technical rules are commonly bent and broken, a practice which may be considered acceptable if the result sounds good. Today's analog standards and practices are the result of over a century of technical evolution, and the origins of many of them have faded into obscurity.

Digital audio is about quantization and data transport, period. It is a deterministic process with rights, wrongs and unforgiving rules that must be followed. Errors in transmission can be checked for, and in some cases corrected by the interface. If there are sonic differences between digital audio formats, they can typically be traced to the analog stages that precede or follow the digital stages. The career path that prepares one for digital audio is university or vocational IT training, or perhaps serious self-study. But ironically, having a computer science degree in no way qualifies one to be an audio practitioner.

It must be remembered that analog is the standard by which digital is judged. Many applications do not require the complexity and sophistication of digital audio. Analog audio isn't going away.

23.10.1 Which Digital Audio "Flavor?"

Many see the future of audio as being digital. But, we know from experience that tried-and-true technologies may move to the background but they never go away. Analog audio is high fidelity, robust, reliable and intuitive. Even "all digital" systems have analog links, such as the connection between a power amplifier and a loudspeaker, or the loudspeaker to the listener. While analog audio is a mature technology that has enjoyed a long development cycle, digital audio is a relatively new development, still in its early evolutionary stages.

The major decision on the part of the system designer regarding digital audio is whether to use a dedicated audio format such as AES3, or an AoE technology. As of this writing the dedicated formats enjoy greater use in traditional sound reinforcement applications that require point-to-point connectivity such as portable systems and auditoriums. AoE offers benefits attractive to very large, multi-purpose buildings and campuses, such as convention centers, airports and casinos. There is no doubt that these distinctions will continue to fade.

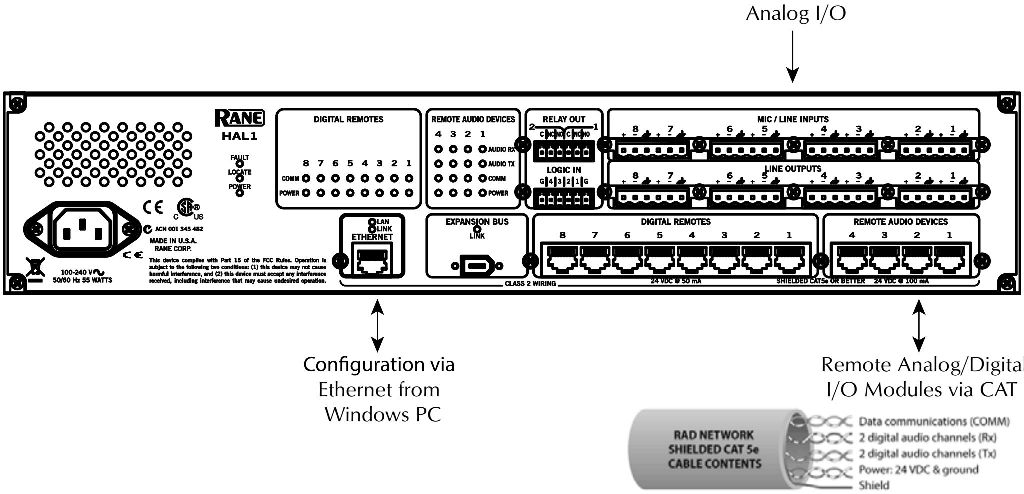

Analog I/O
RANE
HAL1
DIGITAL REMOTES
REMOTE AUDIO DEVICES
RELAY OUT
MIC / LINE INPUTS
LINE OUTPUTS
LOGIC IN
ETHERNET
EXPANSION BUS
CLASS 2 WIRING
Configuration via Ethernet from Windows PC
Remote Analog/Digital I/O Modules via CAT 5
RAD NETWORK SHIELDED CAT 5e CABLE CONTENTS
Data communications (COMM)
2 digital audio channels (Rx)
2 digital audio channels (Tx)
Power: 24 VDC & ground
Shield

Figure 23-20. Hybrid and proprietary approaches can simplify configuration and installation, while retaining sophisticated capabilities. (*Courtesy Rane Corp.*)

23.10.2 Learning Digital Audio

Once the basic theory has been digested, the digital audio learner must get their hands on some equipment. This is where the nuances, quirks, exceptions and requisites are experienced. Most manufacturers of digital audio products provide tutorials, white papers and demo software that cover the specifics of their product line. The time spent reviewing these materials and playing with control software will pay big dividends and help you "connect the dots" regarding how digital audio systems work.

Chapter 24

Sound System Equalization

by Don Davis and Eugene Patronis, Jr.

24.1 System Criteria 555
24.2 Early Research on Equalization 555
    Feedback Defined 556
24.3 The Transient Nature of Acoustic Feedback 556
    Growth 556
    Decay 557
    Early Practitioners 559
24.4 Introduction of Real-Time Analyzers 559
    One-Third of an Octave Analyzers and Equalizers 560
    Forty Years Later 561
    RTA Applications 561
24.5 Band-Rejection, Bandpass, and Band-Boost Filters 562
    Criteria for Band-Rejection Filters 562
    Filter Parameters 562
    Characteristics of Successful Filters 562
    Filter Transfer Characteristics 565
    Minimum-Phase Filters 567
24.6 TEF Analysis in Equalization 567
24.7 How to Approach Equalization 568
    When to Use an Equalizer 568
    Sources of Feedback That Should Not Be Equalized 569
    Feedback Is a Single Frequency 570
    Which Sound Field Is the Microphone In? 570
24.8 What Can an Equalizer Equalize? 570
    Mother Nature's Way 572
24.9 A Real-Time Regenerative-Response Method of Equalizing a Sound System 572
    Where to Put the Microphone for Regenerative Response? 572
    Degree of Correction Necessary 573
    Using Sweep Oscillators 573
24.10 Equalizing for Playback 573
24.11 An Improper Use of Real Time Analysis in Monitoring Music and Speech 574
24.12 Diaphragmatic Absorbers 574
    Room Absorption at Specular Frequencies 574
    House Curves 574
24.13 Don't Equalize for Hearing Loss 575
24.14 Proximity Modes 575
24.15 Checking Microphone Polarity 575
24.16 Loudspeaker Polarity 576
24.17 Summary 576
Bibliography 576

The original one-third octave band rejection filter set utilizing summing circuitry was first used by one of the authors in 1967 and the patent 3,624,298 was filed in March 1969 and issued in November 1971. Since its inception the basic problems in its correct use have been two fold: first, the ability to design a sound system capable of benefiting from the use of an equalizer, and second, the attempt to equalize the unequalizable. These problems are with us more than forty years later.

24.1 System Criteria

Equalization can't solve loudspeaker coverage problems. Equalization can't signal align loudspeakers. Equalizers can't raise acoustic gain unless the system has adequate power available to support the gain increase. Equalizers are of no use in controlling reverberation, discrete echoes, etc.

Careful practice can minimize aggravating these problems via regeneration through the sound system. An equalizer can adjust the direct sound pressure level of the loudspeaker's minimum phase output frequencies. This is accomplished by providing the conjugate amplitude and phase response to any minimum phase aberrations in the loudspeaker's direct sound level, L_D.

Proper equalization adjusts both amplitude and phase to a more uniform response. A delay in microseconds is introduced by the insertion of the filters. The measurement of both amplitude and relative phase is essential in the process of equalization. Group delay for a phono record can be years. The delay through some adaptive digital filters can be appreciable, +30 ms.

The triumvirate of proper equalization, signal synchronization, and seamless coverage is a very powerful tool used to create extraordinary sound quality.

If the system is designed capable of benefiting from equalization, constructed and installed so that coverage is of the proper density, the electrical power is adequate and matched to sufficiently efficient transducers able to absorb it, and the entire system is free from hum, noise, oscillations, and RF interference, you are ready to equalize this system in its acoustic environment to ensure the specified tonal response and acoustic gain at each listener's ears. To do this requires insertion of the necessary filters into the sound system and the taking of meaningful acoustic measurements.

24.2 Early Research on Equalization

Insofar as the authors can discover, the earliest researcher to correctly perform meaningful sound system equalization was Dr. Wayne Rudmose, who at the time of the work to be described was at the Southern Methodist University, Dallas, Texas. Dr. Rudmose published a truly remarkable paper in the journal, *Noise Contro*l (a supplementary journal of the Acoustical Society of America) in July 1958.

We feel the two most authoritative papers ever published on this subject that have retained their fundamental integrity many years later are Dr. Rudmose's and that of one of his students and later an employee of Tracor, Inc. of Austin, Texas, William K. Conner.

Conner's paper "Theoretical and Practical Considerations in the Equalization of Sound Systems" first given at the 1965 AES Convention, appeared in the April 1967, Vol. 15, No. 2 issue of the *Journal of the Audio Engineering Society.* Amid all the nonsense written on this subject, these two papers stand as bedrock for the serious investigator.

In the fall of 1967, one of the authors gave the first paper on a ⅓ octave contiguous equalizer. Wayne Rudmose was the chairman of the session and when the author referred to ⅓ octave as "broad band" Rudmose raised his eyebrow and said "*broad band??*" In 1969, a thorough discussion of acoustic feedback that possessed absolute relevance to "real life" equalization appeared in the *Australian Proceedings of the IREE*. "A Feedback-Mode Analyzer/Suppressor Unit for Auditorium Sound System Stabilization" by J. E. Benson and D. F. Craig, illustrated clearly the step function behavior of the onset and decay of regeneration in sound systems. The aforementioned papers relate the most directly to modern practice and the application of current devices to the adjustment of sound system response. These four sources constitute the genesis of modern sound system equalization.

Rudmose went on to not only introduce ⅓ octave band analyzers, but to describe correctly the cause of "howlback," room resonance effects, cancellations, and the "ringing" encountered in sound systems. In this same paper Rudmose described the devastating effects of transducer misalignments—large holes in the amplitude response and the basic importance of obtaining uniformity of distribution prior to trying to equalize a sound system.

William K. Conner's paper, originally given in 1965, clearly delineated the role of Q and distance on the L_D and $S\bar{a}$ and L_W on L_R. The first thoroughly correct statements regarding useful ratios of L_D/L_R and how to achieve them are presented here. Conner's paper identified the effect of humidity on the sound system's response. Also he pointed out that directional microphones have little to no effect on the power loop gain of the system.

24.2.1 Feedback Defined

Feedback: "*The return to the input of a part of the output of a machine, system, or process.*"

Feedback can be positive or negative in sound systems. Positive feedback, carefully employed, can raise gain. Negative feedback in amplifiers can lower distortion. Hearing yourself speak over a sound system is one form of feedback. It can be beneficial over a monitor loudspeaker or detrimental when the delay causes the talker to stutter. This is a psychoacoustic effect. Some talkers instinctively lower their level upon hearing the monitor; others raise their level, usually a matter of previous exposure.

Oscillatory feedback occurs when the signal from the loudspeaker returns in phase to the open microphone at a level equal to the normal input level resulting in a single frequency "howling" tone. Shock excitation of a sound system on the threshold of sustained feedback may excite many simultaneous tones. The decay of these tones following excitation may be observed on a fast real time analyzer. These tones can then be compensated for one at a time by sequentially introducing equalization at the appropriate frequencies beginning with the tone exhibiting the slowest decay.

The mathematical description given here is a carefully specified single frequency example to illustrate the complexity of the circuitous path. The advent over the past generation of more controlled frequency and polar responses has led to less need of equalizers for the control of feedback. Engineering trade offs, such as controlling directional response at the expense of smooth frequency response, can lead to the legitimate use of an equalizer.

Benson and Craig's detailed explanation of the fundamental mechanism behind acoustic feedback was first published in the March 1969 issue of the *Proceedings of the IRE of Australia.* The paper was entitled "A Feedback-Mode Analyzer/Suppressor Unit for Auditorium Sound System Stabilization." Their analysis made use of the results from an earlier paper by H. S. Antman entitled "Extension to the Theory of Howlback in Reverberant Rooms" that was published in *J. Acoust. Soc. Am.*, Vol. 39, No. 2, February 1966, p. 399 (Letters). A careful study of both of these papers is highly recommended. An abbreviated discussion of the more salient points of these papers related to the transient nature of the acoustic feedback process is presented in the next section.

24.3 The Transient Nature of Acoustic Feedback

Acoustic feedback occurs whenever an open system microphone is exposed to the sound field of the system's loudspeaker array. This means that except for very unusual microphone locations, acoustic feedback is always occurring. What is important, apparently, is not whether acoustic feedback is occurring but rather of the type and degree of the feedback. It is well known that negative or degenerative feedback when judiciously applied in an amplifier can have a desirable stabilizing influence on the amplifier's performance at the expense of reducing the amplifier's overall gain. On the other hand, positive or regenerative feedback applied to an amplifier destabilizes the amplifier's performance, increases its gain, and, if present to a sufficient degree, leads to sustained oscillation. Acoustic feedback in a sound system behaves in a very similar way. The major distinctions between the amplifier case and the sound system case have to do with the feedback path. In the amplifier case the feedback path is usually well defined and the transit time through the feedback path is usually negligibly small. In the sound system case the feedback paths are numerous with appreciable transit times that are dependent upon physical path length.

Consider the simplified situation depicted in Fig. 24-1.

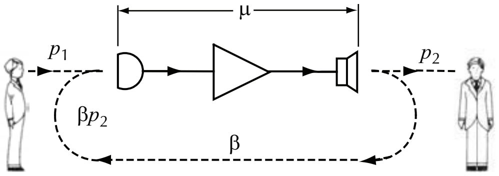


Figure 24-1. Schematic diagram of an elementary sound system with an acoustic feedback path.

24.3.1 Growth

In Fig. 24-1, μ represents the feed forward transfer function of the system. It is made up of the product of the transfer functions of the microphone, amplifier, and loudspeaker. In the absence of any feedback, if p_1 is the acoustic pressure at some reference distance from the microphone, then the acoustic pressure at a similar reference distance from the loudspeaker would be $p_2 = \mu p_1$. In the general case μ would be a complex function of frequency. Similarly, in Fig. 24-1, β represents the feedback transfer function. It, too, is a complex function of the frequency in the general case. As mentioned earlier there are normally many parallel feedback paths, each with their own particular value of β. In our simplified example we will consider only a single feedback path along which there exists a transit time of amount $\Delta t = d/c$ where d is the path length and c is the sound speed. When p_1 is first applied to the microphone, the loudspeaker almost instantaneously produces an acoustic pressure $p_2 = \mu p_1$. After the elapse of an interval of time Δt, a feedback signal of amount $\beta p_2 = \mu\beta p_1$ arrives at the microphone.

Assuming that the original acoustic pressure is still present, the signal at the microphone now becomes $p_1 + \mu\beta p_1 = p_1(1 + \mu\beta)$ and the output now instantaneously jumps to $p_2 = \mu p_1(1 + \mu\beta)$. After the elapse of a second time interval Δt, the input becomes $p_1 + \mu\beta p_1 + (\mu\beta)^2 p_1 = p_1[1 + \mu\beta + (\mu\beta)^2]$ and the output instantaneously becomes $\mu p_1[1 + \mu\beta + (\mu\beta)^2]$. At this point it is safe to generalize the result for an arbitrary number of delay intervals N. After N delay intervals or a total time $N\Delta t$, the output pressure will be given by $\mu p_1[1 + \mu\beta + (\mu\beta)^2 + (\mu\beta)^3 + \ldots + (\mu\beta)^N]$. The analysis so far has placed no restrictions on either μ, β, or on the product $\mu\beta$.

They can each be either real or complex. The remainder of the discussion is greatly eased by requiring only that $|\mu\beta| < 1$, that is, the absolute magnitude of the product of μ with β be less than 1. When this is true the series describing the output converges to the value $\mu p_1[(1 - (\mu\beta)^{N+1})/(1 - \mu\beta)]$. Remember that this output pressure was originally the result of an input pressure signal of p_1. Now if N becomes very large, the term $(\mu\beta)^{N+1}$ becomes vanishingly small and the output divided by the original signal input, system transfer function or gain, becomes $\mu/(1 - \mu\beta)$. This last expression is exactly of the same mathematical form as that for a feedback amplifier. We can simplify the analysis from here on and still obtain meaningful results by letting p_1 be the acoustic pressure associated with a single frequency tone such as 1000 Hz and require that the feedback path is of such length that the return signal is in phase with p_1. This would constitute a case of pure positive feedback. With such a restriction, both μ and β can be taken as real quantities. Let $\mu = 20$ and $\beta = 0.03$. The steady state gain with feedback becomes $20/[1 - (20 \times 0.03)] = 50$. The gain in the absence of feedback is of course μ or in this case 20. The gain ratio is the gain with feedback divided by the gain without or $50/20 = 2.5$. The gain increase brought about by this positive or regenerative acoustic feedback expressed in decibels is $20\log 2.5 = 8$ dB. The effect of this regenerative feedback has been two-fold. Not only has the steady state gain been increased by a really significant amount but also the output arrives at its ultimate value through a series of steps.

24.3.2 Decay

The decay is initiated by the removal of the input p_1. When p_1 is removed, the output drops suddenly by an amount μp_1 while the input continues to be supplied by the feedback component βp_2 that had left the loudspeaker before p_1 had been removed. This continues for a time Δt during which the output remains at the value $\mu\beta p_0$ where now p_0 is the value of the output at the time p_1 was removed. At the end of the interval Δt, the input suddenly drops to $\mu\beta^2 p_0$ thus producing a new output $(\mu\beta)^2 p_0$. This new output persists again for a time Δt. In this fashion the output falls in a series of steps such that after the elapse of N intervals it has become $(\mu\beta)^N p_0$ and the gain ratio has become $(\mu\beta)^N p_0/\mu p_0$.

Fig. 24-2 displays the growth and decay of regenerative acoustic feedback for $\mu = 20$ and $\beta = 0.03$. These values produce a steady state gain ratio of 2.5.

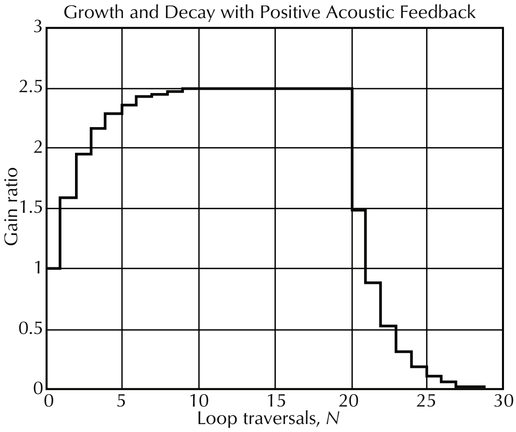


Figure 24-2. Gain ratio versus, N, the number of traversals around the feedback loop.

In Fig. 24-2 the exciting signal is removed at $N = 20$. Note the slow growth and prolonged ringing brought about by the large amount of positive feedback. The system will become completely unstable if $\mu\beta$ becomes 1. In the complete absence of feedback, the gain ratio would step up to 1 at 0 and immediately step down when the exciting signal is removed. Fig. 24-3 shows the situation that exists when the positive feedback is less. In this instance, $\mu = 20$ and $\beta = 0.01$. These values produce a steady state gain ratio of 1.25.

In comparing Figs. 24-2 and 24-3 one can conclude that a larger amount of positive feedback increases the gain ratio and brings about a slower stepwise approach to steady state conditions followed by a longer decay after the removal of the exciting signal. Please see Chapter 14 *Designing for Acoustic Gain*, Section 14.6, The Feedback Stability Margin where the early experimental results of William B. Snow point out the dangers inherent in large amounts of positive acoustic feedback.

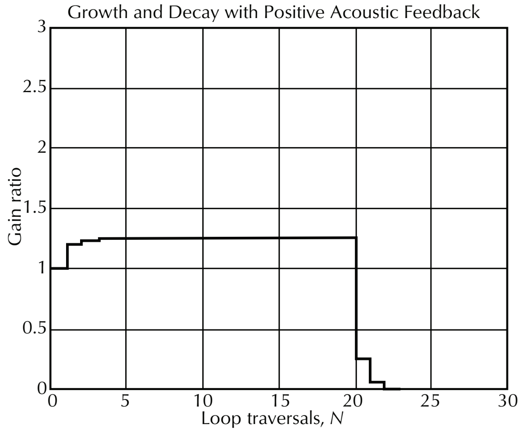


Figure 24-3. Growth and decay with a reduced amount of positive acoustic feedback.

Negative acoustic feedback also has some undesirable acoustic consequences. This is brought about by the relatively long transit time, Δt, around the feedback loop. In a practical case Δt can easily be several milliseconds. Fig. 24-4 illustrates the case where a large amount of negative acoustic feedback is present. In this instance the feedback signal is always of the opposite polarity to that of the exciting signal.

The steady state gain ratio for the conditions of Fig. 24-4 is given by $1/(1 - \mu\beta) = 0.625$. When the initiating signal is turned on at $N = 0$, the gain ratio first overshoots its steady state value of 0.625 and then approaches the steady state value by means of a set of oscillatory steps. At $N = 20$, the exciting signal is removed and the gain ratio now approaches zero through a sequence of constantly diminishing steps. If no feedback had been present, the gain ratio would have stepped immediately up to one at $N = 0$ and would have stepped down immediately to 0 at $N = 20$ when the exciting signal was removed. This

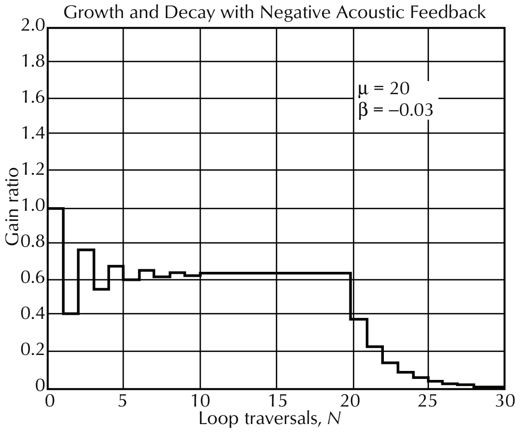


Figure 24-4. Behavior with a large amount of negative acoustic feedback.

situation should be compared with a case of a smaller amount of negative acoustic feedback as illustrated in Fig. 24-5.

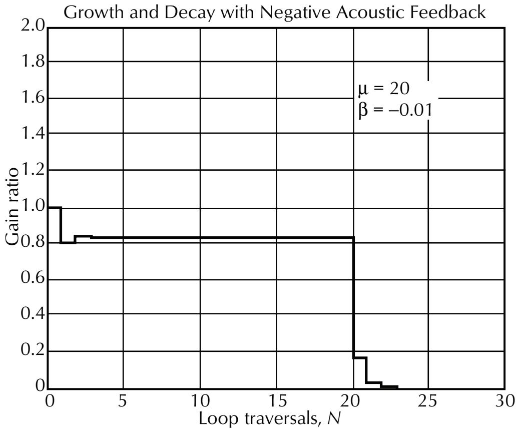


Figure 24-5. The behavior with a small amount of negative acoustic feedback.

In Fig. 24-5, the steady state value of the gain ratio is 0.8333. It should be evident that even negative acoustic feedback distorts the time behavior of the original acoustic signal. Negative acoustic feedback, however, can never produce sustained system oscillation as is true in the positive acoustic feedback case. In the numerical examples we have taken both μ and β to be real numbers just for the sake of simplicity. The general equations are equally valid when they are complex quantities. In such an instance, positive feedback occurs when $|1 - \mu\beta| < 1$ and negative acoustic feedback when $|1 - \mu\beta| > 1$. The case where $|1 - \mu\beta| = 0$ is to be avoided at all costs.

24.3.3 Early Practitioners

Equalization was employed by many early experimenters including Kellogg and Rice in the early 1920s, Volkmann of RCA in the 1930s and most significantly by E. H. Bedell and Iden Kerney's equalizers employed in the 1933–1934 Bell Telephone Laboratories, "Symposium on Wire Transmission of Symphonic Music and its Reproduction in Auditory Perspective."

This remarkable project which transmitted the Philadelphia Orchestra over wired circuits to Constitution Hall in Washington DC with full fidelity and full dynamic range from 35 Hz to 15,000 Hz. Among others it was witnessed by one of my early customers in the 1950s at The Golden Ear hi fi shop. He purchased a full stereophonic Klipschorn system using Ampex 350 tape recorders which he felt came close to the Bell Labs system but did not duplicate the impact he had felt in 1934. William "Cap" Robinson who had served with distinction in World War I and then traveled the world for Chicago Bridge and Iron had had the opportunity to hear all the world's great halls and the majority of the great artists of the first half of the twentieth century.

The sixth paper in this series by Bell telephone laboratories entitled "System Adaptation" had utilized a motor driven oscillator coupled to a level recorder to measure the loudspeakers, the microphone, and their combined effect, as well as learning to account for the temperature and humidity affects on air absorption as well as the influence on the measured curves both nearby the sources and at various points in the reverberant field compromising on a measurement point we'd now call critical distance. They had clearly recognized that the most useful equalization was close to the source, but that due to variability in the distribution of the sound some assessment of the high-frequency rolloff needed to be accounted for out in the reverberant field.

Working with Paul Klipsch in Hope, Arkansas in the 1950s, we had duplicated the work by J. C. Steinberg and W. B. Snow on "Physical Factors," the second paper of the symposium. We traveled to Bell Telephone Laboratories in New Jersey and presented our demonstration to them using two Klipschorns and a middle channel "Heresy" (so named by Mr. Klipsch because it wasn't horn loaded) of two channel and derived third channel geometry allowing the audience to plot their perception of the physical location of the talkers. Unfortunately at that time I was unaware of the additional papers from that symposium.

Dr. C. P. Boner provided a valuable example of repeated application of the principles of equalization by adjusting sound systems that were considered unsatisfactory before equalization into sound systems that satisfied their owners and operators. Often Dr. Boner's work had as much to do with correcting wiring faults, coverage patterns, impedance mismatches, etc., as with equalizing, but in common with all who did such work at that early period, all improvements were attributed to the magic "equalizer."

Until real-time analyzers were available for under $5000 (1970), early equalization work was predominately done by repeatedly raising the gain of the sound system until feedback occurred and adjusting the appropriate filter to reduce the system amplitude at that frequency. One or 2 dB at the very most are necessary to bring the system back to stability. By repeating this for 20–30 feedbacks, it is possible to raise the acoustic gain of a sound-reinforcement system to within a few decibels of unity gain at all frequencies. While this method is an excellent way to increase gain with a nominal amount of test equipment, it leaves the overall tonal balance to the ear of the practitioner. Those with perfect pitch often exhibit a nonuniversal taste, and those with a taste that agrees with the majority of listeners are not often gifted with perfect pitch.

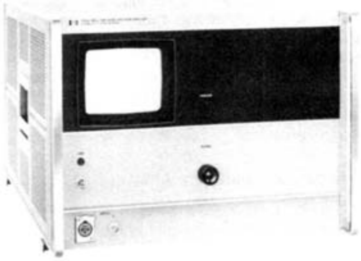
24.4 Introduction of Real-Time Analyzers

Equalization of sound reinforcement systems with the end purpose of increased acoustic gain and enhanced acoustic quality became universal with the introduction of the ⅓ octave equalizer by one of the authors in 1967.

Fig. 24-6 shows one of the authors training a class of sound contractors in January of 1968 in the use of a step-by-step ⅓ octave analyzer for making frequency response measurements for equalization work. The analyzer being used was the GenRad 1564A. In May of 1968 Hewlett Packard delivered to one of the authors the first 8054A ⅓ octave real time analyzer for $10,000 (that's 1960 dollars). The 8054A quickly led to a special stripped down version called the H23-8054A, which came without the nixie tube read out and the frequency selective push buttons, Fig. 24-7. The success of each of the early instruments with the sound contractors who received them led to collaboration with HP which reduced the size and price. Approximately 500 instruments went to the audio industry for use in early equalization work. Other analyzers followed in the 1970s but these were the only ones for several years.

The ability to see, in real time, the result of the equalizer adjustment, in addition to hearing the result, changed the nature of the activity dramatically. The

Figure 24-6. An early equalization training class.



A. H23-8054A

B. 8056

C. 8050A

Figure 24-7. Early RTA equipment.

change from 45 minutes per house curve to 1 second had to be lived through to be truly appreciated.

Today the ⅓ octave analyzer is still one of our most usable audio tools. Heyser-based analysis is more detailed, but there are many cases where the resolution and accuracy of a ⅓, ⅙, and 1/12 octave real time analyzers are ideally suited to the job, such as checking coverage.

The authors feel fortunate indeed to have been the first users in audio to employ and apply ⅓ octave real-time analysis to the study of sound system performance. (See Chapter 11 *Audio and Acoustic Measurements*.)

24.4.1 One-Third of an Octave Analyzers and Equalizers

A common error in the labeling of fractional bandwidth analyzers and equalizers is the term one-third octave. This would imply every third octave rather than one-third of an octave. In actuality they are 1/10 of a decade devices. These devices are widely used in sound systems to analyze the amplitude response of loudspeakers and to smooth that response to allow higher acoustic gain. The best of today's analyzers offer bandwidth equivalents of one-third of an octave, ⅙ of an octave, and 1/12 of an octave. One octave analysis is usually reserved for noise measurements called Noise Criteria curves (NC) plots.

As applied initially in sound system reinforcement work, the goal was the control of acoustic feedback in order to raise the acoustic gain of the system. System resonances, due either to the loudspeaker or aggravated by predominant "room modes", were easily detected by using a sweep frequency oscillator and watching the decay rates of each of the bands on the analyzers screen using the fast scale. The offending frequencies drop at a slower rate than the surrounding frequencies, and by using the sweep oscillator, it was often possible to "narrow in" on the exact frequency causing the problem.

It is theoretically possible to utilize a very narrow band parametric equalizer on such a specific frequency, but the variability of the velocity of sound with temperature can easily cause such a narrow filter to be missed by the frequency being changed by the environment in the room. Because of this problem modern parametric filters offer variable bandwidth as well as amplitude variation.

Quite often the optimum loudspeaker for either efficiency reasons or directivity control reasons is not always the smoothest amplitude response. Laboratory sources usually lacked both sensitivity and directivity, requiring up to 100 times the power needed by the devices actually used. Even in playback system for home use, efficiency and directivity control, while expensive and large, as exemplified by the Klipschorn, easily won out over rigorously engineered "heatsinks."

Another interesting connection between well-engineered loudspeakers and all others is the inexperience of designers to properly understand the role of acoustic radiation resistance. This can provide little to no additional damping in an inefficient acoustic suspension speaker, to a possibly dominant effect in high-efficiency horn systems. The horn system is likely to be very tolerant of low damping ratios, as the cone is coupled to a greater mass of air, so that its own mass becomes less important by comparison. The effect is similar to suppressing the "ringing" of an LC filter by properly terminating the output. All of this comes into sharp focus when reproducing transients and hearing the difference between well-engineered horn systems and consumer products.

In real life balancing economics, smooth amplitude and phase response, directivity control, sensitivity, and physical size, all play a role in a final choice. Pursuing one to the detriment of the others leads to the wide diversity we witness in the marketplace.

24.4.2 Forty Years Later

The authors have each spent over forty years making precision acoustic measurements on sound systems and nearly that long in the practice of equalizing them. The opinions we put forth here are not intended to be final judgments on the matter but rather it is hoped that our suggestions learned from this experience will save you repeating some of the same errors we committed.

Today, the knowledge gained is available and useful in contemplating the next steps. New analyzers are daily revealing the possibilities of manipulating the phase characteristics of a system more directly than in the past, and current digital technology promises full control over the time domain, as well as the frequency domain. As in the past, we expect the future to be shaped by those with the analysis capability coupled to everyday "real life" exposure to sound system design, installation, and operation.

Simple filter networks and their interaction with real life electroacoustic transducers are a complex technical study. To quote Richard C. Heyser,

> *As many technically trained people are prone to do, I naively presumed when I first began analyzing loudspeakers… that I could bring contemporary communication theory to bear and simply overwhelm the poor loudspeaker with technology… I soon found the error in my thinking. The evaluation of the acoustics of loudspeakers and the room containing them proved to represent a microcosm of all the difficult problems in wave propagation. A wavelength range of over 1,000 to 1 is bad enough but the physical extent of the important dimensions in a single experiment range from one thirtieth wavelength to greater than thirty wavelengths for many practical loudspeaker systems.*

The wonder is that the simple tools we employ work as well as they do. It's very fascinating to find that techniques we had to work out intuitively can now be measured accurately and justified as engineering facts.

Equalizers are now omnipresent on the sound system scene. Unfortunately, many are designed as "program equalizers" rather than specifically for the adjustment of systems to maximum acoustic gain. The electronic circuit designer designs to his termination resistor free of all self-doubts or mathematical uneasiness. Then they wonder why some of us are less than enamored of their latest electronic nonsolution to our very real acoustic problems.

24.4.3 RTA Applications

The real-time analyzer has been used by audio engineers for almost 40 years and has proved invaluable in the following areas:

1. House curves made with a measuring microphone.
2. House curves made with a sound system microphone.
3. Examining distribution of all frequencies at differing locations at the same time.
4. Examining house curve at the performer's location.
5. Response curves of the filter settings.
6. Detecting feedback frequencies.
7. Frequency response of microphones to be used in the sound system.
8. Examining crosstalk between lines.
9. Setting levels throughout sound system areas both electrically and acoustically.
10. Detecting resonating surface areas by observing the effect of manual damping of the vibrating surface.

Pink noise (equal energy per octave) is used rather than white noise (equal energy per hertz) because the bandpass filters used in the typical RTA are constant percentage bandwidth rather than constant bandwidth. This means that a white-noise signal put into a constant-percentage bandwidth analyzer would have a +3 dB/octave rise with increasing frequency. The filters grow wider with increasing frequency, thereby summing more energy at the same level. Pink noise on a per hertz basis decreases 3 dB/octave (10 dB/decade); therefore, pink noise matches constant-percentage bandwidth response, allowing a flat response across the screen of an RTA. See Section 24.11, *An Improper Use of Real-Time Analysis in Monitoring Music and Speech*.

Fig. 24-8 shows an example of how unavailable acoustic gain, due to feedback caused by highest amplitude present, can be made available by equalizing all frequencies, making them equal in amplitude response.

24.5 Band-Rejection, Bandpass, and Band-Boost Filters

All types have been used to equalize sound systems. The author's preference is for band-rejection filters. Boosting a narrow band of frequencies is not a natural acoustic phenomenon. Any two frequencies can come together in a room to a maximum (in the practical case) of +3 dB but may combine to complete cancellation. The only thing "narrow band" going on in an acoustic environment is rejection, never summing, though acoustic focusing may be mistaken for summing. Those who think there are narrow acoustic "peaks" in the environment are the same ones who advocate equalizing low frequency room modes, i.e., usually by trying to boost the null frequency. Very narrow peaked responses can and do lurk in the transducers used in sound systems. Such devices should be replaced, not "equalized." Again, we have often witnessed skilled electronic circuit designers trying to correct a deep notch in the frequency response with a very sophisticated electronic filter when the cause of the notch is one driver out of alignment with another.

Interesting to end users, but a matter of disgust to creative circuit designers, the simpler, almost sloppy, filters seem to work the best.

24.5.1 Criteria for Band-Rejection Filters

So far as the authors are concerned, if the following criteria are met, it should be a satisfactory filter:

1. Minimum phase response.
2. Combining, sometimes called summing.
3. Minimum excess delay.
4. Not narrower than ⅙ octave bandwidth with ⅓ octave preferred.
5. Band rejection-type with maximum depth of 14 dB, preferably in 1 dB steps.

This is not to say that there are not interesting differences in designers' choice of circuits. Fig. 24-9 shows two exceptional fine passive equalizers wherein the only observable difference is in the unterminated state. Though largely not manufactured today, these passive units never wear out. It would be a pity to discard them when newer electronic equipment is installed, but one needs to be aware of the need of the "buildout" and "termination" requirements.

24.5.2 Filter Parameters

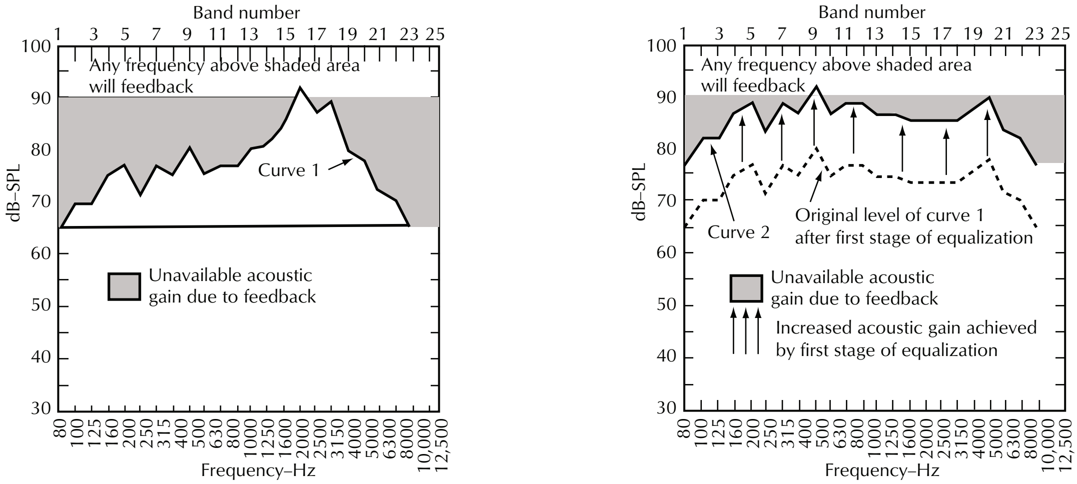
Figs. 24-10A and 24-10B show the parameters of a very narrow-band filter for both amplitude and phase. The meaning of the bandwidth of a filter is shown in Fig. 24-11.

One of the earliest methods of sound-system-room equalization employed individual broadband networks to shape a rough inverse of the house curve, Fig. 24-12A. This was followed by the insertion—one at a time—of very narrow notch filters at the predominant feedback frequencies, Fig. 24-12B. When the two sets of filters were electrically combined, the response was like Fig. 24-12C. Fig. 24-12D is the replacement tuning finally put in the job after the advent of ⅓ octave-spaced bridged-T filter sets. Note how the same overall gain restoration is provided by either type of filter, but that the ⅓ octave-spaced filter set removed feedback at many frequencies by changing the slope rate instead of depressing the amplitude. It is important to remember that the sound system deviations from uniformity in a system worth equalization are quite correctable by the slope-rate changes available in the ⅓ octave-spaced filter sets.

24.5.3 Characteristics of Successful Filters

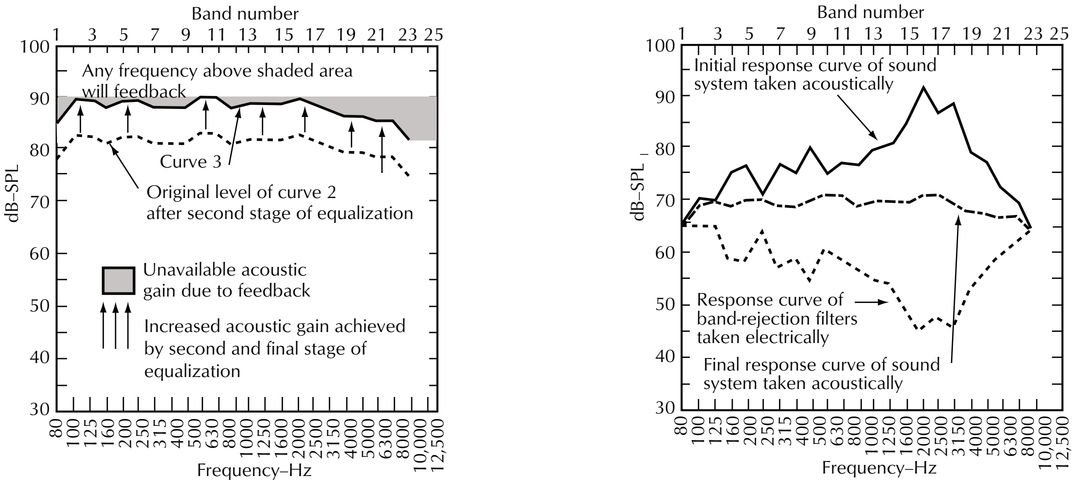
Let us compare an ISO ⅓ octave bandpass filter with a ⅓ octave spaced band-rejection filter, Fig. 24-13. By further comparing the two basic bandpass filter shapes, both ⅓ octave and ⅒ octave, we can see that the inverse of the ⅒ octave response most closely approximates the response of the band-rejection filter, Fig. 24-14. Looking at a single filter section and recording each of its 1 dB steps, we get a series of curves as shown in Fig. 24-15.

If we were to record each filter section in a set of 24 such filters, by setting each filter for a maximum rejection and then restoring it to zero before recording the next filter section set at its maximum rejection, we would obtain the set of curves shown in Fig. 24-16. If we were to turn all 24 filter sections



A. Typical situation that can occur in an auditorium. Curve 1 shows the sound pressure output of the loudspeaker if a signal equal in level at all frequencies is connected to the input of the sound system. The irregularity of the output is partly due to the inability of the loudspeaker to respond perfectly uniformly to a uniform input signal, and mostly due to the effect of the room itself on the acoustic output of the loudspeaker. Feedback will occur first at 2000 Hz, since any attempts to raise the gain will find this peak to be the first frequency to push above the limit line. To follow this example, assume the acoustic gain is 10 dB.

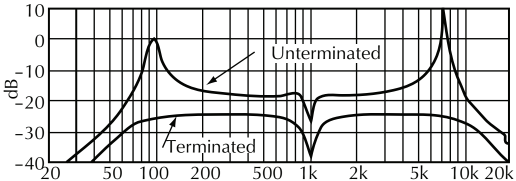
B. Curve 1, after the peaked area between 2000 Hz and 3150 Hz, has been equalized to the majority of the other frequencies. The arrows indicate how all these frequencies may now be raised simultaneously in gain before the new peak of 400 Hz pushes above the hatch lines and causes feedback. The number of decibels at each frequency between curve 1 and curve 2 respresents the increased acoustic gain made possible by the first stage of sound system equalization.



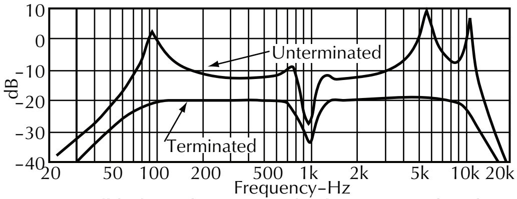
C. Additional smoothing of the curve can allow greater acoustic gain at the majority of frequencies. However, further smoothing, even if perfectly done, would yield either 1 dB or 2 dB at the very most throughout the frequency region of critical importance for speech. By comparing curve 2 with curve 3 we can see that through the vital frequency response area for speech, for example, the acoustic gain at all frequencies is increased from 300 to 3000 by, typically, 10 dB or more. Originally, only 2000 Hz could be brought to 90 dB Lp before feedback occurred; now all frequencies can be brought to 90 dB L_p before feedback.

D. The electrical response curve of the critical band rejection filters (bottom curve) join together to form the inverse of the loudspeaker room response curve (top curve). Because this inverse filter response curve is included in the total sound response, the smoothed over-all acoustic response shown by the middle curve results.

Figure 24-8. How equalization raises acoustic gain.

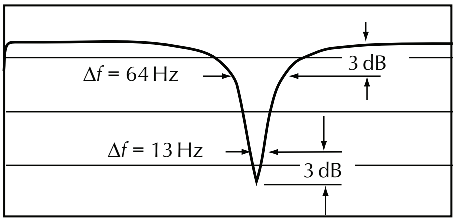


A. Well behaved passive in both terminated and unterminated states.

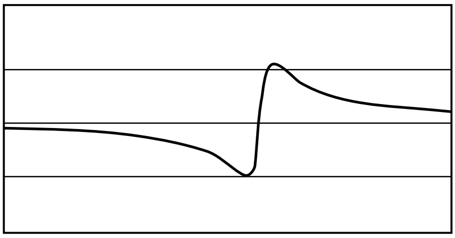


B. Well behaved passive in both terminated and unterminated states—note difference between high frequency behavior of this equalizer compared to that in Fig. 24-9A.

Figure 24-9. Frequency response curves for a passive equalizer (1000 Hz filter set at −14 dB with high pass frequency at 80 Hz and low pass frequency at 10 kHz.)



A. Amplitude

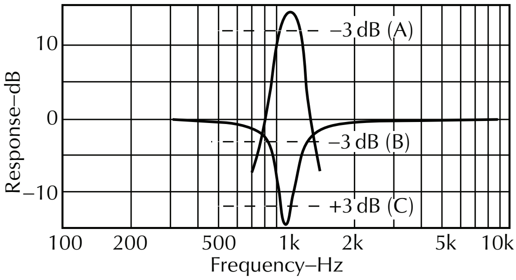


B. Phase

Figure 24-10. A very narrow-band filter.

to maximum rejection at the same time, we would discover their most important property—they combine, Fig. 24-17. Note that when they combine they are essentially additive (their combined depth exceeds 20 dB at the bottom of the ripple).

Fig. 24-18 shows in detail how they combine. The narrower curve is 1000 Hz set at –6 dB. The wider curve is 800 Hz at –2 dB, 1000 Hz at –2 dB, and 1250 Hz at –2 dB. Here they have combined to become 6 dB deep, and the center of the curve is at the middle of the three. It is not difficult to imagine the complexity of combining from 14 to 24 of these sections all at different levels to appreciate that some form of real-time observation is required to



1. The bandwidth of an active bandpass filter is measured 3 dB below its center-frequency level (A).
2. The bandwidth of a band-rejection filter is measured 3 dB below the normal level before the filter is inserted (B).
3. On occasion individuals have chosen to define the bandwidth of a band-rejection filter as "up 3 dB from the center notch."

Figure 24-11. The bandwidth of an active bandpass filter.

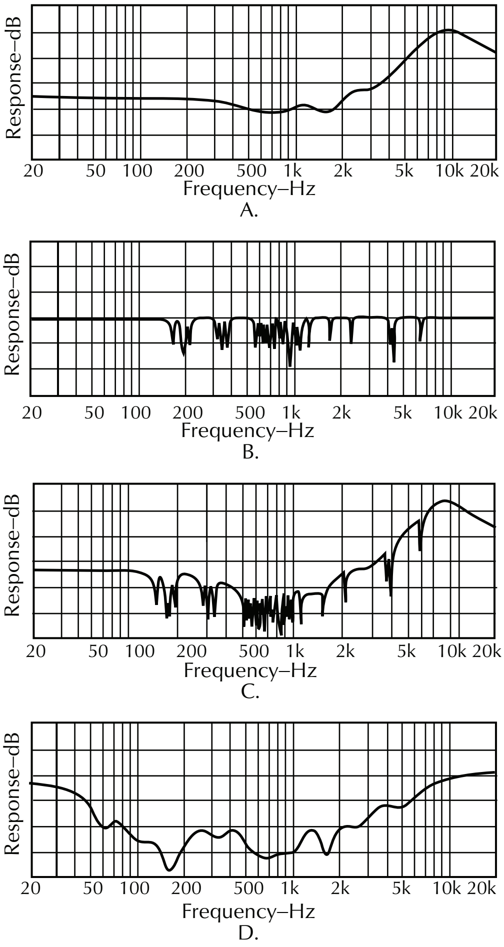


Figure 24-12. Very narrow band equalization (A, B, and C) and typical combining type filter equalization of the same system (D).

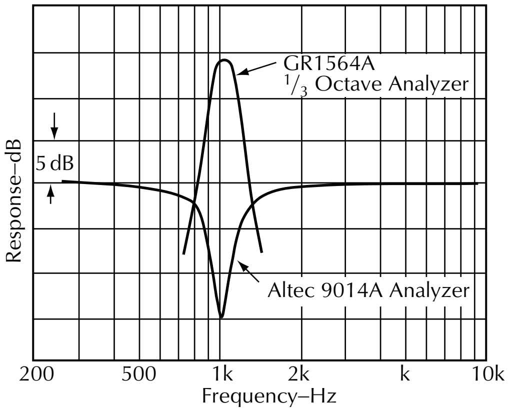


Figure 24-13. Broad-band combining-type band-rejection filter section compared to ⅓ octave active bandpass filter section.

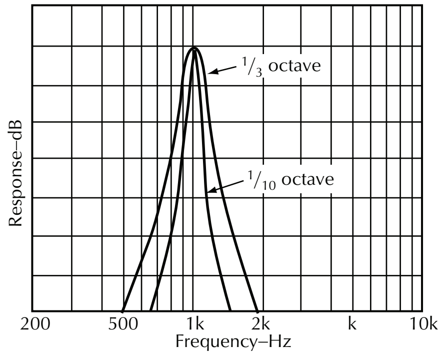


Figure 24-14. Comparison of a ⅓ octave bandpass filter with a ⅒ octave bandpass filter.

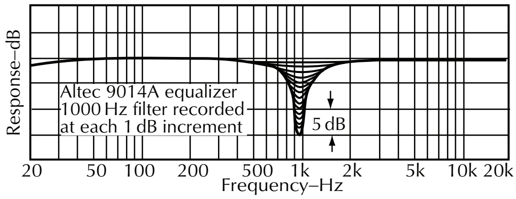


Figure 24-15. Response of a band rejection filter at each of its 14 steps.

comprehend thoroughly what is going on. Through such combining, the smoothest conjugate phase response is achieved. Fig. 24-19 shows the electrical response of a set of filters on an actual job. Fourteen sections were employed at the frequencies and levels indicated. Note that at 160 Hz the filter is at –10 dB but the curve is at –18.5 dB due to combination effects. The real test is to compare the inverse of the filter response before equalization. This is done on a ⅓ octave basis in Fig. 24-20. Filters producing the type of response we have just looked at can be either passive or active.

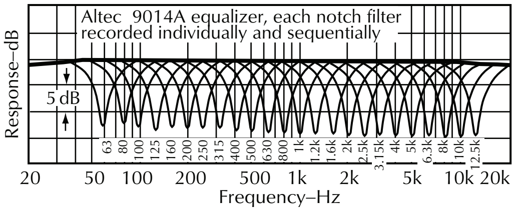


Figure 24-16. A series of band-rejection sections recorded sequentially one at a time.

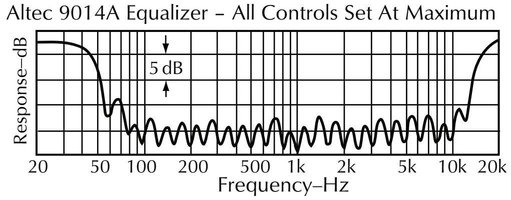


Figure 24-17. Series of band rejection sections all simultaneously turned to maximum attenuation.

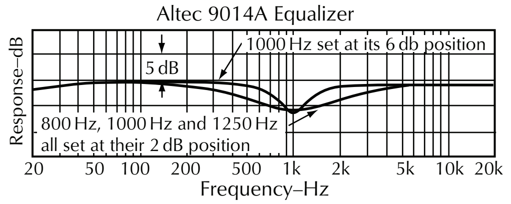


Figure 24-18. Response of a 1000 Hz section with 6 dB attenuation compared with the response of a combination of 800 Hz, 1000 Hz, and 1250 Hz sections with 2 dB attenuation each.

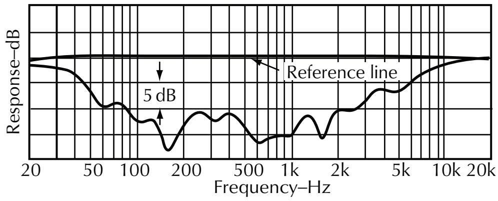


Figure 24-19. Example of band filter section in combination to form inverse of raw house curve.

Fig. 24-21 compares the electrical amplitude response of a set of very narrow-band filters and a set of critical-bandwidth combining filters after adjustment on the same job.

24.5.4 Filter Transfer Characteristics

The transfer characteristics of band-rejection, minimum-phase filters are shown in Fig. 24-22. In the previous example of the combining power of these filters, in terms of amplitude, we see an example of their combining power in terms of

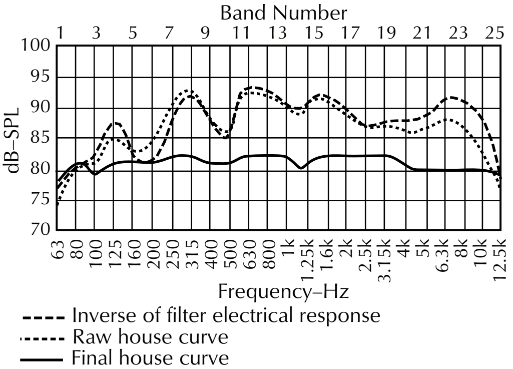


Figure 24-20. Example of correlation between raw house curve and inverse of filter electrical response curve.

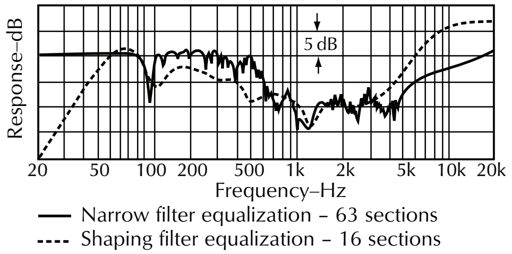


Figure 24-21. Comparison of response of a set of very narrow-band filters and a set of critical-bandwidth combining filters.

phase, Fig. 24-23. The steeper the amplitude slope rate, the steeper the rate of phase change.

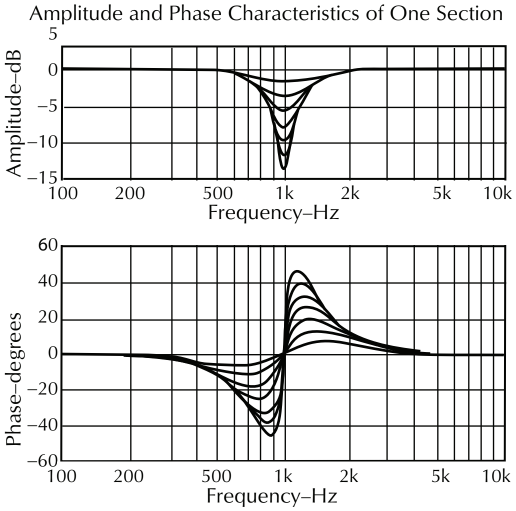


Figure 24-22. Bridged-T configuration for tandem filter sections.

Fig. 24-24 shows all the filters at –1 dB amplitude and the resultant phase characteristic. Fig. 24-25 shows the same information for –4 dB.

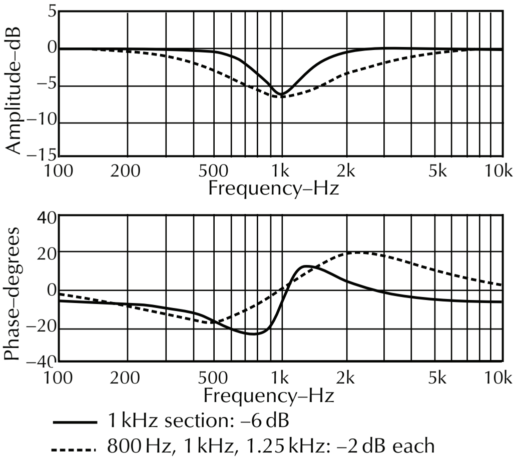


Figure 24-23. Active configuration for tandem filter sections.

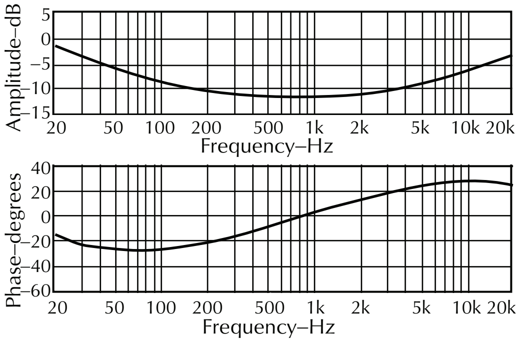


Figure 24-24. Combined phase and amplitude response, all sections set for –1 dB.

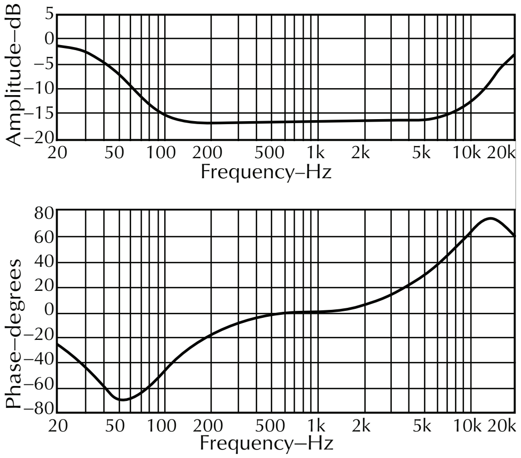
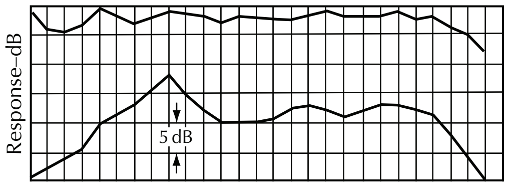


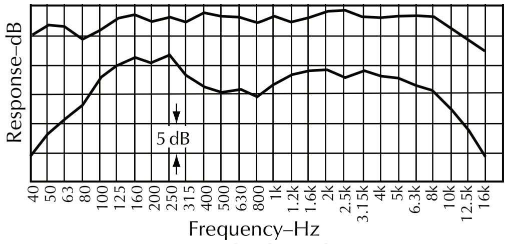
Figure 24-25. Combined phase and amplitude response, all sections set for –4 dB.

Taking a practical example, Fig. 24-26A shows the equalized and unequalized response of the left channel of a monitor system. Fig. 24-26B shows the same information for the right channel. Fig. 24-27A shows the electrical amplitude response of the corrective filters, and Fig. 24-27B shows the phase response of each channel.

if

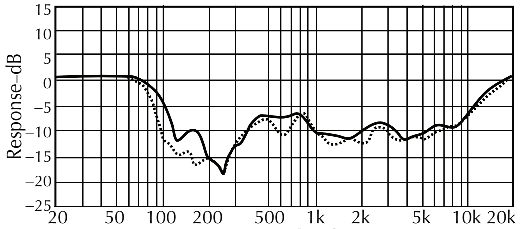


A. Left channel.



B. Right channel.

Figure 24-26. Equalized and unequalized response of a monitor system.



A. Amplitude

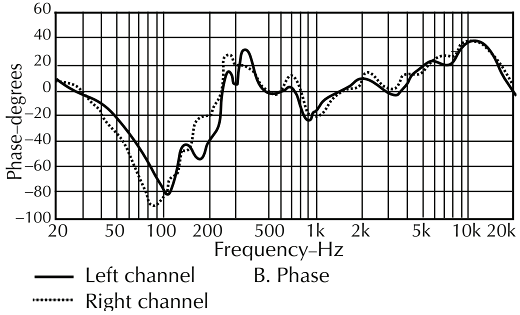


Figure 24-27. Response of equalizers.

24.5.5 Minimum-Phase Filters

A minimum-phase filter introduces the minimum possible phase shift but still retains the corrective amplitude change. It is also obvious that the relative phase between channels is virtually identical. It is this careful band-by-band resolution of phase that so many listeners have dubbed as the "sharp focus" that equalization seems to produce in sound systems already relatively smooth in an amplitude sense.

Richard C. Heyser has pointed out that:

> *Highly important to a loudspeaker's ability to produce accurate sound, when it has been properly equalized, is that of minimum phase change. A minimum-phase-change loudspeaker is one in which, when all amplitude response variations are removed by conventional resistance, capacitance, and inductance networks, it has the minimum possible phase shift over the frequency spectrum. Properly designed equalizers for balancing the amplitude response* ***will also automatically balance the phase response for a minimum phase loudspeaker.*** (Bold added.)

Also:

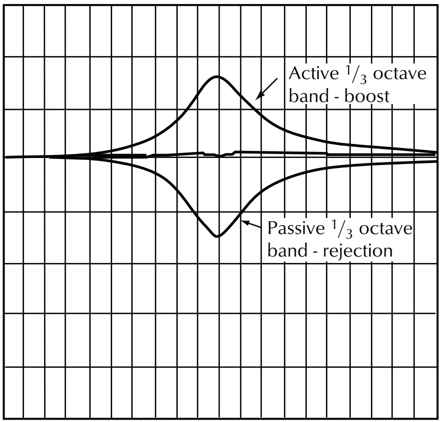
> *A nonminimum-phase loudspeaker will usually exhibit frequency-response difficulties which can be associated with signal delay effects which, in turn* ***cannot*** *be corrected with conventional passive or active equalization.* (Bold added.)

Fig. 24-28A shows an active ⅓ octave band-boost filter (BBF) raised 9 dB, a passive ⅓ octave band-rejection filter (BRF) lowered 9 dB, and the resultant smooth amplitude response. Fig. 24-28B shows the phase response of the BBF and the BRF as well as the resultant smooth phase response.

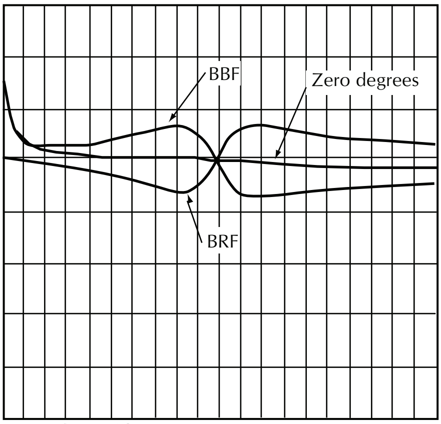
The distinction between a bandpass filter (BPF) and a band-boost filter (BBF) is that the "skirts" of a BPF continue on to minus infinity (–∞). The skirts on a BBF return to zero reference after passing through the peak. Such filters are found in active parametric filter sets and extreme caution is advised in their use in live system stability adjustment. One must exercise caution when it comes to program effects' adjustment.

24.6 TEF Analysis in Equalization

With TEF analysis in use for almost all serious equalization work today, the engineer can view L_D, L_R, and L_{RE} (early reflection levels), separately or together as desired. The level, direction, and time of travel for each reflective interference can be observed. An aberration in L_T can be segregated and if it is caused by L_D equalized, if caused by L_{RE} blocked, and if due to L_R, it can be treated in the



Vertical: 6 dB/division
Horizontal: Auto 0.00–2000.24 Hz
Resolution: 1.0010E + 01 Hz
A. EFC



Vertical: 45°/division
Horizontal: Auto 0.00–2000.24 Hz
Resolution: 1.0010E + 01 Hz
B. Phase frequency curve, PFC

Figure 24-28. Notch filter in opposition to boost filter.

statistical manner. Formerly well hidden transducer aberrations, particularly in phase and in time behavior, are now strikingly evident and consequently easily prevented.

As a consequence of this enhanced ability to actually "see" what's going in the total system, both rapidly and accurately, the design procedures are modified to incorporate this new knowledge and we are finding that less and less equalization is required in the newer systems. Where and when equalization is needed, it is invaluable. Misapplied, it can create harsh sounding high frequency distortion, instability in the system, and worst of all, a belief that equalization has solved a problem that in actual fact is still unaddressed.

Fig. 24-29A is the Envelope Time Curve, ETC, of a packaged system with a separation between the low frequency unit and the high frequency of 0.17 ft, 0.15 ms. (See Chapter 11 *Audio and Acoustic Measurements* for a full explanation of the Envelope Time Curve.) Fig. 24-29B is the magnitude and phase of the same loudspeaker with the phase response made as smooth as possible. Fig. 24-29C is the Nyquist display of the same loudspeaker with the cursor set at a high frequency that has encircled the origin of the display indicating a *non-minimum phase* system. Failure of a cursor in a Nyquist plot to rotate clockwise as frequency increases indicates a mis-selected signal arrival time at the measurement microphone. Harry Nyquist of the Bell Telephone Labs in the old days was truly a genius.

24.7 How to Approach Equalization

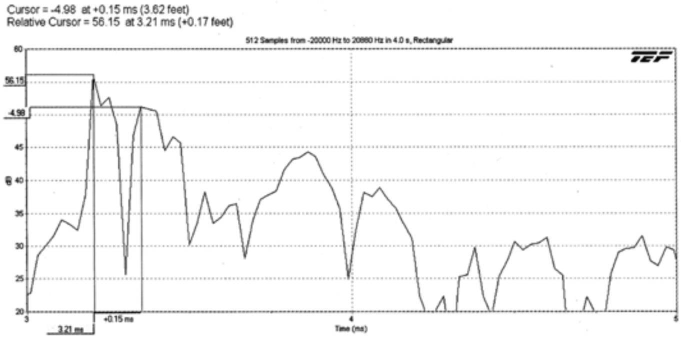
Gently! Slowly! These are key words for key attitudes when utilizing equalizers. After every adjustment, listen carefully to the remaining sounds. The goal is to improve sound quality in sound reproduction systems as well as increase acoustic gain in the case of reinforcement systems. In any type of system, stop tuning and examine the system with care when the equalizer causes a detrimental change in the sound quality.

In using equalizers, we can borrow from the medical fraternity and say, "First, do no harm."

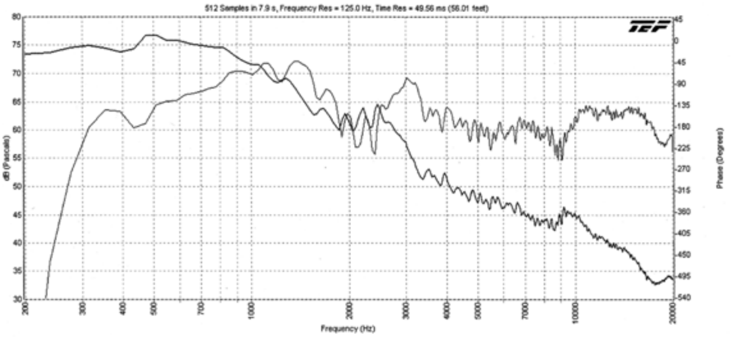
Anything within a system that requires steep slope rates (over 18 dB/octave) or excessive amplitude change (in excess of 3 dB) in order to control a given frequency increment (on the order of ⅓ of an octave) should be subject to serious consideration as to its replacement. In today's marketplace there is a sufficient number of very well-behaved electronics, microphones, loudspeakers, and interconnection networks available to allow avoiding the use of inferior products. Increasing understanding of the Heyser Transform and its ability to allow us to understand the transformations in time at given frequencies coupled to the newer analyzers to measure such parameters should quickly lead to both better components and more skillful application and adjustment of them.

24.7.1 When to Use an Equalizer

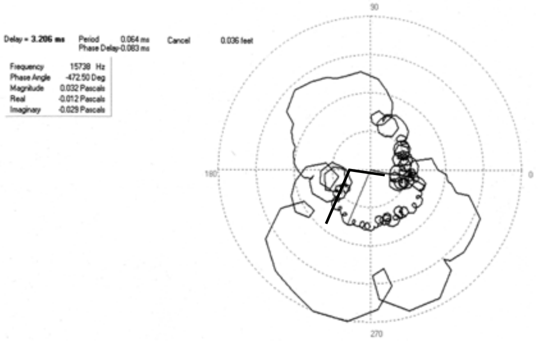
Equalization, in common with most sound system components, can be misapplied. Therefore, a discussion of when and where it is appropriate to use an equalizer and when and where it should not be employed is useful.



A. Envelope time curve of a packaged two way loudspeaker system arrival times and relative distances plus levels.



B. Magnitude and phase of a packaged two way loudspeaker system.



C. Magnitude is length of the cursor at any frequency. The angle of the cursor is the phase. Encirclement of the origin reveals non minimum phase behavior.

Figure 24-29. Envelope Time Curve, magnitude, and phase with the phase response made as smooth as possible and the Nyquist display of a loudspeaker.

Let's first imagine a very special case wherein we have a loudspeaker that already has the acoustic response we desire and its coverage pattern exactly covers our audience area so that each listener is receiving the identical level and $\%AL_{CONS}$. Let's further assume that no sound reflects off this audience into the reverberant space. In such an ideal case, we would require no equalization unless for some reason we desired program equalization or the deliberate distortion of program material for some departure from this ideal case. This clearly identifies the fact that only some departure from this ideal case might require correction. A deviation in L_D might then require an equalizer. Other deviations would require different remedies.

24.7.2 Sources of Feedback That Should Not Be Equalized

Equalizers are most misused by end users in the following areas:

1. Used to correct instabilities caused by comb filters designed into the system rather than removing the cause of the comb filter.

2. A microphone on a desk stand near a hard surface.
3. Excessive insertion loss per filter.
4. Use of the filter to control a problem caused by mechanical feedback.
5. Use of the filter to control feedback caused by crosstalk between circuits.
6. Use of the equalizer to adjust the steady-state response of devices whose transient response is an undamped resonance.

There are many very critical areas where the distinction between acoustic gain equalizers and program equalizers becomes quite hazy. Our increasing appreciation of right brain-left brain influences in the perception of the received signal by the listener causes us to proceed with caution in interpreting the technical data inundating users of analyzers. Perhaps we should separate electroacoustic transducer magnitude and phase adjustment from other forms of signal processing that are quickly coming on the scene (e.g., signal synchronization).

If we were to reserve the word “equalizer” for those devices that made system gain equal at all frequencies and used the words “signal processor” or “adjuster” for those devices intended to shape specifically the transfer function of the system for any purpose other than optimum acoustic gain, we might forestall at least a little of the confusion.

24.7.3 Feedback Is a Single Frequency

Acoustic feedback is indeed single frequency. The fallacy that is usually being defended by that statement is the implication that since feedback is single frequency, so should the compensating filter be single frequency.

Nyquist, Waterfall, and Antman have clearly shown the amplitude, phase, and signal delay path causes of feedback. A single frequency feedback is often less than 0.1 dB greater amplitude at that frequency than the surrounding frequencies.

The only extremely narrow band effects the authors have ever observed in real life sound systems is band rejection such as phase cancellation or diaphragmatic absorption. Recall that the acoustic environment is passive, not active.

24.7.4 Which Sound Field Is the Microphone In?

Fig. 24-30 hints at the multiplicity of sound fields that might be encountered in a single acoustic environment in connection with the operation of a sound reinforcement system. Being alert to each of these acoustic fields can solve many problems that seem mysterious when considered in the context of a single field. Often, you will equalize in more than one of the fields at the same time. For example, you will equalize the main system for the audience, the foldback system for the entertainer, a delayed under-balcony system for a distributed system, and separate equalization for a signal-delayed portion of the main system. Some of these are in the direct field and some in the reverberant field. No one, as yet, has all the answers to applying equalization to the fantastic variety of sound systems being designed today.

Fig. 24-31A illustrates the time dependency of each of the sound fields. Fig. 24-31B depicts the frequency dependency of the characteristics of a sound field. Finally, Fig. 24-31C shows the level dependency of sound fields relative to their distance from the sound source.

In almost all sound system measurements, we avoid the “near field” of an acoustic source. Most measuring microphones used with equalization measurements are placed in the far reverberant sound field. Equalization of monitor loudspeakers in recording control rooms almost always are in the far free field.

When using modern analysis capable of separating L_D from L_R, the operator is able to choose between direct sound levels, L_D, early reflected sound levels, L_{RE}, reverberant sound levels, L_R, and total sound levels, L_T (L_T is comprised of L_D, L_{RE}, L_R, and L_N, where L_N is the ambient noise sound field).

One of the most frequent errors made in using ⅓ octave analyzers, or FFTs which measure L_T, is the failure to note the level of L_N separately followed by observation of its effect on L_T.

In complex multi-loudspeaker arrays the best practice is to turn on one loudspeaker at a time for equalization. If two loudspeakers share a common coverage area they are first looked at individually and then adjusted combined.

24.8 What Can an Equalizer Equalize?

The question, “What can an equalizer equalize?” needs to be asked. Some claim to equalize the room. Is this possible? We think not. When an electronic or passive equalizer is installed in between a mixer and a power amplifier we need to know that all we can equalize is the electrical signal being sent to the loudspeaker.

What can an audience do to affect L_D from a sound source? The answer, of course, is absolutely nothing. Therefore, it is clear that the audience can

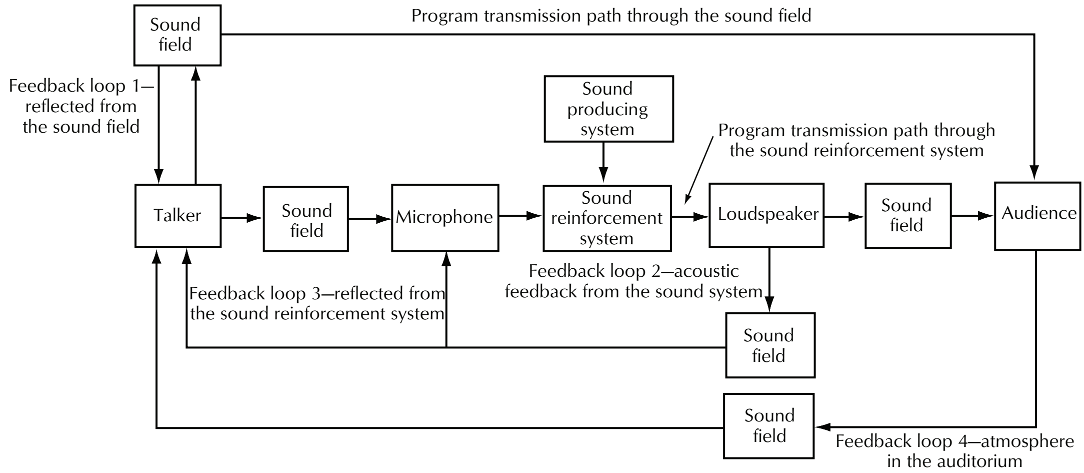
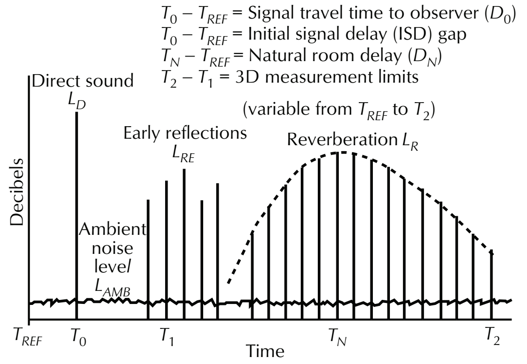
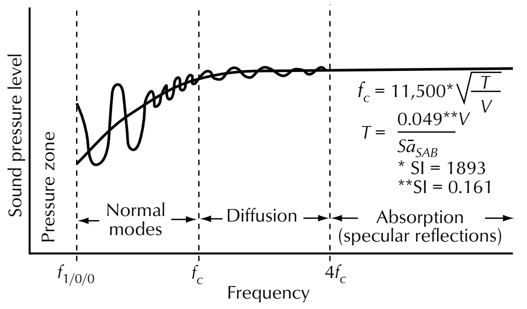

Sound field
Program transmission path through the sound field
Feedback loop 1—reflected from the sound field
Sound producing system
Program transmission path through the sound reinforcement system
Talker
Sound field
Microphone
Sound reinforcement system
Loudspeaker
Sound field
Audience
Feedback loop 2—acoustic feedback from the sound system
Feedback loop 3—reflected from the sound reinforcement system
Sound field
Sound field
Feedback loop 4—atmosphere in the auditorium

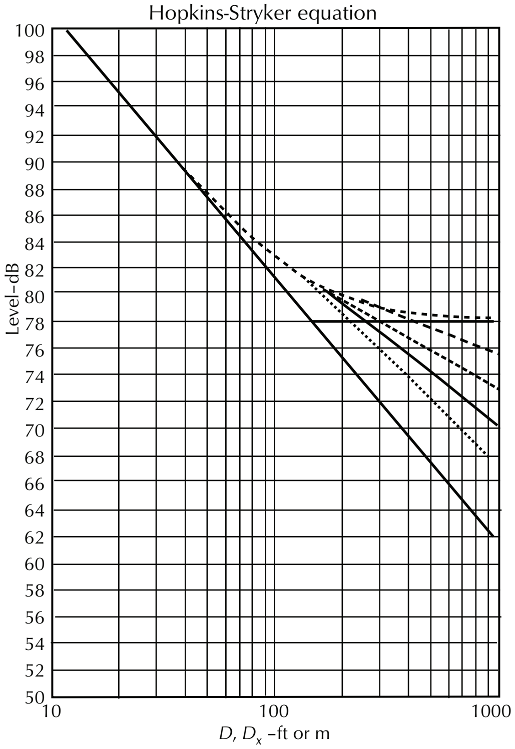
Figure 24-30. Example of the multiplicity of sound fields that can affect how you choose to equalize.



A. Large room acoustic response.



B. Controllers of steady state room acoustic response.



C. Acoustic level versus distance.

Figure 24-31. Acoustic level versus distance and time.

only alter L_{RE}, L_R, and L_N. Now, ask yourself the question, "how can an equalizer adjust L_{RE}, L_R, and L_N?" The answer is that it cannot. I would hesitate to mention such obvious facts except for the remarkable number of times in the popular press that claim the contrary.

If one has only an RTA, one needs to equalize the loudspeaker's L_D. Often, due to the location of the loudspeaker; the microphone is in the reverberant sound field. Knowledge of the free field response of the loudspeaker allows for intelligent "guesses" as to what aberrations in the reverberant sound field should not be responded to with the equalizer, i.e., focused reflections.

If the facility is already built and functions are taking place, be sure to attend a "performance" before installing the equalization and follow up with a visit at a performance after equalization. It is important to test the equalization with speech. Often only music will be played which gives little idea of speech clarity and often a separate equalization is required for speech and music.

In one case, whenever the equalization was made uniform to the desired house curve, there was not enough acoustic gain in the audience area. When the acoustic gain was raised by the feedback method, the shape of the house curve in the audience area was unacceptable. Analysis with an RTA revealed that the loudspeaker array mounted in the proscenium area was not properly shock mounted and was causing the arch structure to reradiate a signal downward to the microphones, causing premature feedback. When this large array was properly shock mounted and properly isolated in the proscenium arch area, then the house curve could be shaped as desired, and the acoustic gain could be brought to its potential at the same time.

Another cause of the same effect is when the rear of the proscenium is open to the stage house and the rear of the array has substantial radiation of its own into the microphone. The house curve in the audience may be a good response, yet the microphone on stage has severe feedback problems because it is receiving a strong radiation from the loudspeaker, usually in the bass region.

24.8.1 Mother Nature's Way

There are in audio and acoustics both natural and unnatural distortions. An example of a natural distortion is harmonic distortion because we hear harmonics in nature. An example of an unnatural distortion is print through on tape recordings where we hear the echo first followed by the desired sound. This never occurs in nature, thus our brain is extremely sensitive to its occurrence. So it is with "boost filters."

In nature any two signals can combine and go to various depths depending upon the relative phase angle between them. Under ideal conditions, they can only add to 3 dB greater levels and psychologically we notice the presence of something far more than its absence. Subjectively, over more than forty years of system equalization in the field, qualified investigators report slope rates in excess of 18 dB/octave as audible. Therefore, filter sets (equalizers) should be designed to avoid steep slope rates, be combining, and not introduce unnatural distortions (boosting).

24.9 A Real-Time Regenerative-Response Method of Equalizing a Sound System

One of the earliest demonstrations of the effect of operation near regeneration on the measured frequency response of a sound reinforcement system was by William B. Snow before the February 1954 meeting of the Audio Engineering Society in Los Angeles. Snow recorded on a high-speed graphic level recorder the dramatic amplitude changes that occurred in the overall amplitude-versus-frequency response as the reinforcement system was brought nearer and nearer regeneration. (See Chapter 14 *Designing for Acoustic Gain*.)

This same method is still used with a manually operated oscillator to identify those frequencies whose amplitude responded unduly to the approach of the regeneration point of the sound system. Shock excitation was employed to observe the increased decay periods of frequencies that otherwise would not feed back upon being increased in gain but were unduly affected by the approach of the system regeneration point. Snow's paper had also demonstrated this point, proving that such frequencies, when shock excited near regeneration, could take as much as 4 to 6 times as long to decay as the same frequencies required when they were excited well below the regeneration point (−12 dB).

Over 40 years of active participation in equalization of sound systems has shown us that making regenerative response curves is one of the most useful techniques applicable to equalizer adjustment. More often than not, the frequencies that ring as they are swept do not come up into steady-state-feedback yet they interfere with speech intelligibility and the overall acoustic gain of the system. The regenerative response curve technique allows careful analysis of both the electrical and the acoustic responses while allowing both the regenerative and degenerative frequencies to be identified.

24.9.1 Where to Put the Microphone for Regenerative Response?

Where to place the microphone for viewing the equalization with an analyzer is one question. The second question is where do I place the microphone that is to cause the regeneration of the signal in the sound system?

One good practice is to place a measuring microphone out in the main coverage pattern looking at the array. Then place the regenerative microphone (the microphone that is to be used to cause the system to go into acoustic feedback during the test) in the location where it is to be employed for normal system usage.

24.9.2 Degree of Correction Necessary

Interconnect the instruments and the sound system as illustrated in Fig. 24-32. The instrument labeled "analyzer receive" can be a ⅓ octave of ⅟₁ octave real-time analyzer, a TEF analyzer, a high-quality wave analyzer, or a suitable FFT. The instrument labeled "analyzer send" can be a logarithmic sweep oscillator (for use with the constant percentage bandwidth analyzers), a linear sweep oscillator (for use with the TEF or FFT), or a random noise generator. In fact, it can be just about any controlled source that is capable of exciting the sound system's entire bandpass.

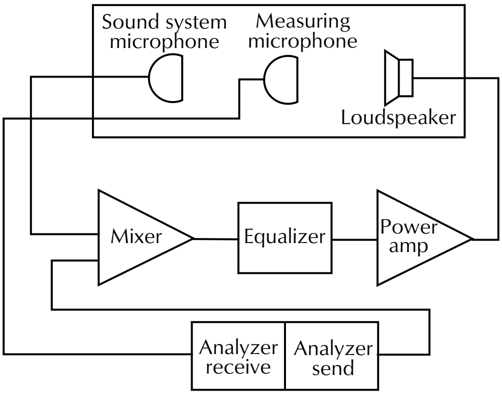


Figure 24-32. Regenerative response tuning.

Clearly evident in this process is the fact that below regeneration the amplitudes of the frequencies involved are of the order of 1 or 2 dB higher than other nearby frequencies. As these frequencies approach regeneration, their amplitudes can "swell" to +20 dB or more. (See Chapter 14 *Designing for Acoustic Gain*, Figs. 9-7 and 9-8.) Naturally the sound system must supply the power for this regeneration, as the room cannot—it is passive. The room can provide nulls and peaks, but not amplification. If, before regeneration is approached, one of these frequencies has its amplitude brought to uniformity with the remainder, then upon approaching regeneration again, this frequency will not "swell" in amplitude.

By using sweeps from an oscillator, it is then possible to watch what happens to the sound system's electrical and acoustic responses as well as to listen to the transient response. (Measure the transient response effects if a TEF analyzer is available.)

The ability to make the needed controlling corrections at a magnitude associated with the nonregenerative state of the system was first observed by the authors in 1966 and was presented as a paper.[*]

24.9.3 Using Sweep Oscillators

Rapid sweep oscillators allow very effective regenerative response tuning. By rapid rate we mean a full sweep from 20 Hz to 20,000 Hz at about one-half the time it takes mid-range sound to decay 60 dB in the environment where the tuning takes place. Typically, sweep rates of 10 kHz/s are useful. If searching for a low frequency anomaly, use a logarithmic sweep. If searching for a high frequency anomaly, use a linear sweep.

This rapid sweep will cause all the bands on a ⅓ octave equalizer to jump up on the screen and then drop at the integration rate of the analyzer (usually 0.1 s at the "fast" rate). A "ringing" band will drop slower than the rest of the bands being shock excited by the sweep. The sound created by the sweep passing through the system and shock exciting the nonlinear frequencies has a "gong-like" tone not unlike that heard in older department stores to summon or alert personnel. As the equalizer's amplitude is adjusted slowly while listening to the "gonging," it turns into a gong-in-a-pillow sound, indicating sufficient attenuation is present at that frequency.

Pink noise for regenerative excitation is more useful at lower frequencies. Bringing the system to within a few dB of regeneration using pink noise allows the room and sound system to "display" which frequencies are unduly sensitive to approaching regeneration including those due to phase as well as amplitude.

24.10 Equalizing for Playback

The real time regenerative response method can be startlingly effective in the equalization of sound reproduction systems or sound synthesizing systems. In this case, instead of using a performer's microphone to achieve regeneration, the calibrated measuring microphone is simultaneously used for both room-response measurement and regeneration. Multichannel systems tuned using this technique are characterized by superior spatial geometry as well as improved tonal response. The improved reproduction of geometry is believed to be due to the better acoustic phase response between channels at the listener's position, and while only a small area so benefits, it is usually only the mixer's general area that has to be covered in the typical studio monitoring room situation. Some unusually extended-range systems may now feed back at frequencies well above audibility (in one case above 30 kHz), and care must be taken to use a low-pass filter in conjunction with either the sound system or the measuring system. Of course, attention must be paid to avoid tuning in the null of a standing-wave pattern. A short walk with the microphone of the analyzer, especially in a small control room, is fascinating, instructive, and necessary.

---

[*]. D. Davis. "Adjustable ⅓ Octave Band Notch Equalizer for Minimizing Detrimental Interaction Between a Sound System and Its Acoustic Environment," presented at SMPTE Meeting in Chicago (Sept. 1967) and AES Meeting in New York (Oct. 1967).

24.11 An Improper Use of Real Time Analysis in Monitoring Music and Speech

Constant percentage bandwidth filters have absolute widths that increase in direct proportion to the center frequency of the filter. When performing spectrum analysis with instruments based on such filters it is necessary to employ a random noise source whose spectrum has constant energy per octave, i.e., pink noise as opposed to a noise source that has constant energy per unit bandwidth, i.e., white noise.

A system possessing a uniform or flat response on a per unit bandwidth basis that is excited with pink noise will produce a flat display on a constant percentage bandwidth analyzer. Such a system excited with white noise would produce a response that rises at 3 dB/octave on a constant percentage bandwidth analyzer. Therefore, when constant percentage bandwidth analyzers are employed to study the spectra of program material where it is desired to determine the response displayed on a per unit bandwidth basis, it is necessary to precede such an analyzer by a filter that has a response that falls at the rate of 3 dB/octave. Any evaluation of program material without such a device is invalid.

It is the authors' belief that this uncorrected error is why so many professional mixing engineers still use meters and indicators in place of the much more useful real time analyzer. Trained ears didn't agree with the uncorrected visual display. The noise control people made their criteria constant percentage bandwidth based, thereby judging relative results. The recording engineers, home hi-fi enthusiasts, and other researchers did not realize the need and therefore failed to compensate for it.

24.12 Diaphragmatic Absorbers

Care should be observed in the handling of dips in the response of a loudspeaker and a room caused by diaphragmatic action of some boundary surface. This is identifiable when, after all the bands around the dips are brought down, they still fall the same number of dB below the surrounding bands. Do not chase it on down, because that will only increase the insertion loss of the total equalization with but negligible improvement in the response. The correct method is to drive the loudspeaker room combination with a tunable bandpass filter and observe the effect on the real-time analyzer to find the frequency where the absorption of the signal is greatest; then use your fingertips and feel all the surfaces of the space, including walls, doors, windows, etc. You will feel the offending surface vibrating in sympathy with the test signal, in one case a walled-up window area.

At a famous recording studio during a demonstration of equalizing monitor loudspeakers, a diaphragmatic absorption was traced to a loose "sound-lock" door. Upon holding the door tightly shut, an 80 Hz notch in the house curve disappeared.

24.12.1 Room Absorption at Specular Frequencies

It has been common practice for the past fifty years to adjust the high frequency response of sound systems to compensate for high frequency absorption in the room. Analysis suggests that the loss of high frequencies being compensated for does not occur as a result of the absorption present, but as a result of L_W lowering rapidly at about the same frequency Q increases, with the resultant illusion that the response is uniform, but duller. The most common cause of radical high frequency loss in sound systems is either device misalignment or a high level, very early reflection (i.e., within 2 ms or less).

24.12.2 House Curves

A famous acoustician once was heard to say that the "house curve" (i.e., the response as viewed on a ⅓ octave real time analyzer) should be down 10 dB at 10 kHz referenced to 1 kHz. What his listeners forgot to consider was where he was standing. It was 70 ft to 80 ft out in front of a horn type loudspeaker system. At that distance, when you take into account air absorption, microphone diffraction characteristics, and high-frequency distortion components, 10 dB down seems quite sensible. When you are in a control room 10 ft from a loudspeaker the 2 to 3 dB typical of microphone diffraction at 10 kHz is more logical and air absorption is not a factor. When necessary to err, then err on the side of a little extra rolloff.

24.13 Don't Equalize for Hearing Loss

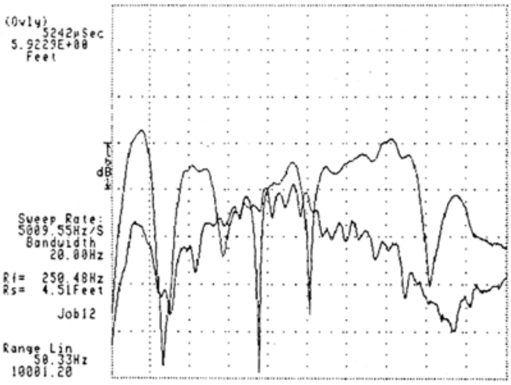
Many times there is a tendency to attempt to adjust the amplitude response of a sound system to make it the inverse of the hearing-loss curve. This is not a good idea for several reasons:

1. Young people with normal hearing will be annoyed.
2. Older people have made mental compensation for the gradual onset of the loss and would also be annoyed. They usually desire the overall level higher.
3. Available high-frequency drivers would have their distortion increased noticeably with such a boost.

24.14 Proximity Modes

The microphone proximity effect, traditionally referred to in the technical literature, is the effect of increased bass response in the microphone as the talker gets closer to the unit. This remains true of most unidirectional microphones today and is often effectively used by trained performers to enhance their otherwise weak bass tones. Since the advent of sound system equalization, however, we have become aware of still another effect of the proximity of large bodies (performers) on a typical cardioid microphone. That is the increased tendency to feedback at some key midrange frequency where the system is otherwise stable until the microphone is approached. You can use your hands cupped around the microphone to bring the system into feedback and can adjust the level of feedback by "playing" the microphone. In adjusting the appropriate filter, care should be taken not to carry the adjustment too far. The idea is to correct the tendency of the microphone to cause instability when it is approached by the performer and not to remove all tendency toward feedback even when the microphone is completely encircled by a closed hand. TEF analysis has shown that this instability is caused by "comb filters" produced by reflected sound from the performer.

One classic example was Dan Seals who is a left-handed guitar player and was having trouble with acoustic feedback whenever he turned to his left at the microphone. He allowed us to make a measurement with the same setup as when he was performing. Fig. 24-33 shows our measurement. The guitar reflection and the hat brim reflection combined acoustically at the microphone to cause a genuine excess gain problem. When he turned to the left, the body of his guitar reflected the left monitor towards the microphone and his hat brim reflected the right monitor to the same place. When Dan Seals saw the measurement, he said, "*We have met the enemy and they is us*!"



Upper trace is Dan Seals with cowboy hat and guitar
Lower trace is open microphone

Figure 24-33. Effects of surfaces on feedback.

24.15 Checking Microphone Polarity

Surprisingly, one minor checkout prior to equalization time that often is overlooked is the poling of the microphones in a multimicrophone system. The old way was to arbitrarily assume that the first microphone you picked up was correctly poled. Holding it in one hand and the second microphone in the other hand, and bringing them closer and closer together while talking into them (such as "*one-one-one*"). They were in polarity if the apparent bass response increased as they were brought closer together in front of your mouth. They were out of polarity if the bass weakened as they were brought together. In one memorable case the "first" microphone was reversed and this simple process reversed all the others. The arrival on the scene of a polarity checker revealed the error. In any case, be sure to check this important factor before equalizing. A polarity reverser is invaluable in this work. Today we know to check for absolute polarity as it has been repeatedly demonstrated that it is audible on speech. TEF analyzer phase measurements instantly indicate the correct polarity (as well as, in the TEF case, the difference between polarity and phase). Be careful before you rewire microphones; the patch cords could be miswired.

24.16 Loudspeaker Polarity

In examining the "raw" response of a loudspeaker array, pick the poling that gives the most usable response through the crossover region. True phasing can enter in here, as well as polarity, and great care should be exercised in the relative positioning of horns to each other, especially the spacing and positioning of the high-frequency elements in relation to the low-frequency elements. Remember, out in the audience area there will be phase relationships between direct and reflected sound as well as those between two direct sound sources. The real-time analyzer is invaluable for examining the potential variations and their effects on the audience area. Today, through TEF analysis we have identified signal misalignments of from fractions of an inch to about one foot as particularly hazardous to speech quality.

24.17 Summary

The advent of practical sound system equalization in situ in the late 1960s coincided with the development of portable ⅓ octave constant percentage bandwidth real-time analyzers which led to a revolution in the design, installation, and operation of sound reinforcement systems.

The availability of equalizers and analyzers quickly led to the training of large numbers of alert sound contractors, consultants, and operators in sound system measurements. Proper design led to much more powerful loudspeaker arrays constructed by those knowledgeable about directivity factor, comb filter interference, and signal delay and synchronization. Manufacturers responded with vastly improved data.

Today we have unimagined design aids, loudspeaker data, and a cadre of knowledgeable users. Today equalization is a small component in the cornucopia of tools available, but it does have the satisfaction of having been the catalyst to dramatic improvement in the design and installation of outstanding sound systems.

Bibliography

J. E. Benson and D. F, Craig. "A Feedback-Mode Analyser/Suppresser Unit for Auditorium Sound-System Stabilisation," *Proc. IREE, Australia* (Mar. 1969).

W. K. Connor. "Theoretical and Practical Considerations in the Equalization of Sound Systems," *Audio Eng. Soc.*, Vol. 15 (Apr. 1967).

D. Davis. "Facts and Fallacies on Detailed Sound System Equalization," *Audio* (1969).

_______. "A Real Time Regenerative Response Method of Equalizing a Sound System." AES Paper (May 1973).

Gene Patronis. "Elementary System Theory," *Syn-Aud-Con Tech Topic*, Vol. 21 No. 3 (Spring 1994).

W. Rudmose. "Equalization of Sound Systems," *Noise Contr.*, Vol. 24 (July 1958).

M. R. Schroeder. "Improvement of Acoustic-Feedback Stability by Frequency Shifting," *J. Acoust. Soc. Am.*, Vol. 36 (Sept. 1964).

B. Snow. "Frequency Characteristics of a Sound-Reinforcing System," *J. Audio Eng. Soc.*, Vol. 3 (Apr. 1955).

G. Stanley. "Minimum Phase: Defined and Illustrated," *Syn-Aud-Con Tech Topic*, Vol. 5, No. 10 (1978).

Chapter 25

Putting It All Together

by Eugene Patronis, Jr.

25.1 Acoustical Analysis 579
25.2 Alternative Solutions for a Given Space 579
  More Modern Treatment 580
  Virtual Sound Processor 581
25.3 Device Interconnections 582
25.4 Analog Interconnection Circuitry Types 583
  Balanced Circuits 583
  Balanced Circuits and Susceptibility to EMI 584
  Brief Description of Electromagnetic Waves 584
  Shielding 586
  Unbalanced Circuits 587
  Cables, Connectors, and the Pin 1 Problem 588
  The Pin 1 Problem 588
  Removal of SCIN from Unbalanced Circuits 589
  Consumer Output to Balanced Input 589
25.5 Signal Cables—Analog Audio, Digital Audio, and Video 590
  Video and Digital Audio Signal Cables 590
  The Cable Problem 591
  Characteristic Impedance 592
  Attenuation 592
  Phase Velocity 592
  Skin Effect 593
  Measurement of Cable Signal Propagation Properties 595
  Sample Calculation 596
  To Terminate or Not to Terminate 596
25.6 AES3 597
  Encoding Format 597
  General Transmission Characteristics 598
  Electrical Characteristics 599
  AES Information Documents and Unbalanced Transmission of AES3 600
  Sony/Phillips Digital Interface Format 600
  Microphones with Digital Outputs 601
  Connectors for Digitally Interfaced Microphones 602
25.7 Computer Control and Communication of Digital Audio 602
  A Little History 602
  Networking of Digital Audio Data 603
  CobraNet® 604
Bibliography 604

The diverse nature and broad range of configurations of devices that constitute sound and sound reinforcement systems is impressive indeed. A sound system designer may be called upon to design a system as simple as a paging-background music system for a retail venue. Alternatively, the requirement could equally as well be the design of a system accommodating the needs of a multi-purpose arena having a seating capacity greater than the population of a small city. A competent designer would treat both of these with the same degree of professionalism. The responsible designer, in addition to formulating a design that accomplishes the necessary acoustical goals, must also tailor the design to match the capabilities and training level of the personnel who will be called upon to operate and maintain the system. The writer was taught as a youngster that a good hunter must always adjust the caliber of the weapon to match the game being sought. One does not pursue Cape buffalo with an air rifle nor does one hunt squirrels with 155 mm howitzers. Overly complicated systems can and do fall in disarray if the operating personnel lack sufficient expertise. Under-designed systems often fail to meet the required acoustical and operational goals. The design of a system for a given venue and purpose, unlike the solution to certain math equations, is not unique. There are more ways than one to skin a cat and some cat skinners are more artful than are others.

25.1 Acoustical Analysis

System design always begins with an acoustical analysis of the space involved. Many of the preceding chapters have been devoted to the various aspects of acoustics that ultimately bear on the determination of loudspeaker properties and loudspeaker arrangements necessary to provide sound having adequate coverage, intelligibility, and level. Once the required loudspeaker properties are at hand a search of manufacturers' specification sheets can be initiated to identify the particular loudspeakers that may be employed in the system. Acoustical analysis also determines the type or types of loudspeaker arrangements required, in particular whether the requirement is for a single source array, such an array supplemented by satellite arrays, or a distributed set of loudspeakers, etc. Once the loudspeaker arrangement is decided upon it becomes possible to determine the power amplifier requirements. It is obvious at this point that we are working backwards from the loudspeakers toward the input of the system in building up the design. Rather than proceeding further with a general discussion it may be more informative to work through some alternative designs for a particular acoustical space while making use of the full range of audio technology that is presently available.

25.2 Alternative Solutions for a Given Space

The space to be considered is that of a large reverberant house of worship having a fan shaped seating plan. In order to achieve speech intelligibility in this space while simultaneously maintaining source identity it is necessary to employ a central source supported by step delayed satellite sources. A plan view of the space appears in Fig. 25-1.

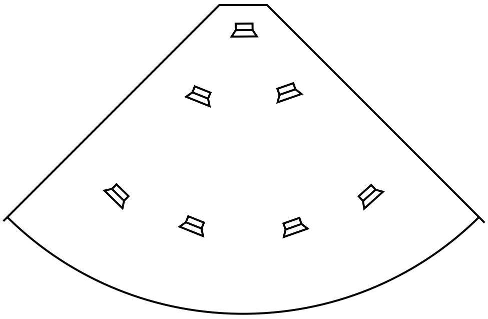


Figure 25-1. Plan view of seating space illustrating loudspeaker placement.

The loudspeakers are each elevated in a similar manner above the seating areas and tilted downward. The area covered by each loudspeaker is approximately the same. This is accomplished by positioning the satellite loudspeakers on two circular arcs centered on the central source and increasing the loudspeaker density proportional to the square of the arc radius. In this instance only two arcs are required necessitating two steps of signal delay. The loudspeakers themselves are full range, three way units, and in the initial treatment are considered to have appropriate passive crossover networks. Each loudspeaker is allotted 200 W. The single line diagram for the initial treatment is presented in Fig. 25-2.

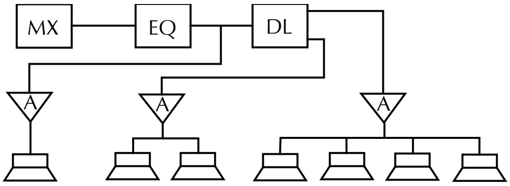


Figure 25-2. Single line diagram for basic treatment.

The basic treatment is essentially all analog with the single exception being the two step signal delay. Even the signal delay has an analog input and two analog outputs. For the sake of simplicity attenuation or voltage sensitivity settings are not shown. A single equalizer is sufficient, as all loudspeakers are the same and mounted similarly. The equalizer can be all active or passive with a gain makeup stage. All circuitry is balanced in and out with the exception of the power amplifiers. The power amplifiers have balanced inputs and unbalanced outputs. From left to right, the power amplifiers are progressively 200 W, 400 W, and 800 W. One variation on this basic treatment worth consideration is that of having individual power amplifiers for each loudspeaker. This variation appears in Fig. 25-3.

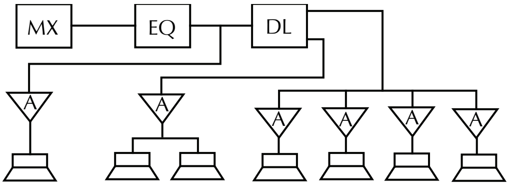


Figure 25-3. A variation on the basic treatment.

In the variation of Fig. 25-3 all of the power amplifiers are the same and some redundancy has been introduced. If a spare amplifier is provided for the source loudspeaker, the system can limp along when a power amplifier fails. In both the basic treatment as well as in the variant, the system can survive failure of the equalizer. Dedicated equalizers are provided with bypass switches and should be operated as unity gain devices. Bypassing a failed equalizer will produce a change in sound quality but not in sound level.

The basic treatment as well as its variant employing a digital signal delay would have first been a possibility in the early 1970s as digital signal delays were not available prior to that time. Prior to that time signal delay had to be accomplished by analog only techniques that are crude as compared with present day techniques. Short delays were accomplished by having a loudspeaker drive a plane wave tube that was provided with pickup microphones spaced at appropriate distances along the tube. Longer delays could be achieved with modified tape recorders provided with a single recording head and several physically spaced reproduce heads. The plane wave tubes had to be physically longer than the maximum delay distance interval in order to allow for a non-reflective termination and had to be isolated in long attics or other such spaces. The tape recorders had to employ continuous running high-speed tape loops that were prone to breakage without warning. The "good old days" were not always that good in certain respects.

25.2.1 More Modern Treatment

A more modern treatment of the acoustical space discussed above relies more heavily on digital signal processing circuitry. Following along with this change the three way loudspeakers will be crossed over actively rather than passively with the crossover function preceding power amplification. The single line diagram for this approach appears in Fig. 25-4.

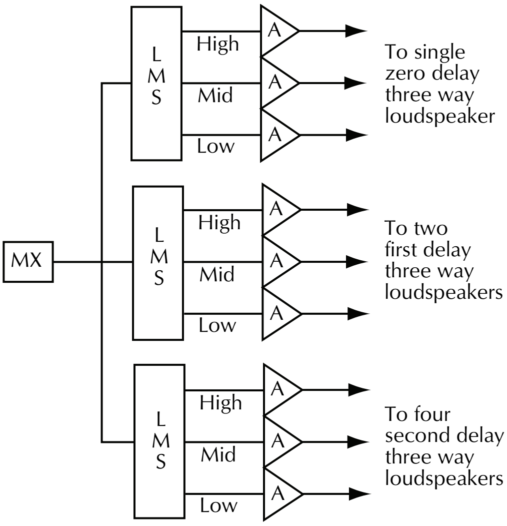

MX
L M S
High
Mid
Low
A
To single zero delay three way loudspeaker
L M S
High
Mid
Low
A
To two first delay three way loudspeakers
L M S
High
Mid
Low
A
To four second delay three way loudspeakers

Figure 25-4. Treatment employing increased digital signal processing.

In the treatment of Fig. 25-4 the mixer can be all analog or a digital mixer with either analog outputs or digital outputs featuring AES/EBU connectivity. This treatment introduces loudspeaker management systems or LMS. These units when first introduced featured only balanced analog inputs and outputs. They have subsequently evolved to the point where both analog and digital inputs and outputs are available. The digital inputs and outputs conform to the AES/EBU standard. All internal processing is performed by DSPs. An LMS can incorporate all of the following functions:

1. One-third octave equalizer.
2. Parametric equalizer.
3. Signal delay.
4. Choice of crossover type and order.
5. Compressor/limiter.

The units are usually configured to offer four analog/digital inputs with either four or eight analog/digital outputs. Setup may be accomplished with front panel controls aided by a display screen or through a serial interface such as RS/EIA 232 to a computer provided with the appropriate software. The computer in turn may be part of a larger network. The employment of these units with a digital mixer is particularly desirable for a number of reasons other than the option of direct digital connectivity. Even though emphasis has been placed on the main house system in the stated example, houses of worship also feature many auxiliary systems. Typically there are separate listening systems for the choir and for the hard of hearing. Additionally, the mixer must provide separate audio feeds for radio broadcast, television broadcast, and recording. Digital mixers have the capability to store several different preset scenes that may be called up according to the dictates of the particular program being presented. Digital mixers in addition to the ordinary analog inputs also accept direct digital inputs such as AES/EBU, SPDIF, and MIDI.

A variant on the treatment described immediately above replaces the conventional analog in, analog out power amplifiers with amplifiers that accept a digital input. These amplifiers are based on conventional class D or similar amplifiers that have been modified to convert pulse code modulated signals into the pulse width modulated signals required by amplifiers that feature switching in the output stage. With such amplifiers, the required digital to analog conversion occurs in the low pass filter following the switching output stage. At this point, analog signals exist only at the system input and output transducers with all required signal processing having been accomplished in the digital domain.

25.2.2 Virtual Sound Processor

LMS is well adapted for many installations but does have limitations. Many hardware boxes are required for large installations and the list of functions performed by an individual unit falls short of what may be required in many systems.

Peavey Electronics Corporation addressed these shortcomings in 1993 through the introduction of the MediaMatrix® system. This system is designed around a software based, integrated sound system design, control, and hardware platform. The hardware platform utilizes a modular computer mainframe with a variety of supporting disk drives, a system controller board, and several digital signal processing boards. The system features a graphical user interface supplied with a library of several hundred software audio devices. The scope of this library is such that the user may fashion practically any conceivable audio system. The user simply selects the required devices from the library and "wires" them up on the computer generated display and control surface.

Analog signals are received from or communicated to the outside world via supporting hardware analog break out boxes. Similarly, digital signals are received from or communicated to the outside world via supporting hardware CobraNet® break out boxes. Real time network communication and control of digital audio signals is by means of 10/100Base T Ethernet accompanied by CobraNet® hardware interface. In order to produce a complete system, the user need only supply the original audio signal source or sources in either analog or digital form and the necessary power amplifiers and speakers.

The MediaMatrix® system constituted the first viable virtual sound system processor. This system also featured internal diagnostics and the capability to expand to almost any size required. It can be employed in our simple example system as illustrated in Fig. 25-5 but more importantly, it equally as well can be employed to serve all of the audio needs of a giant international airport system as well as any system in between.

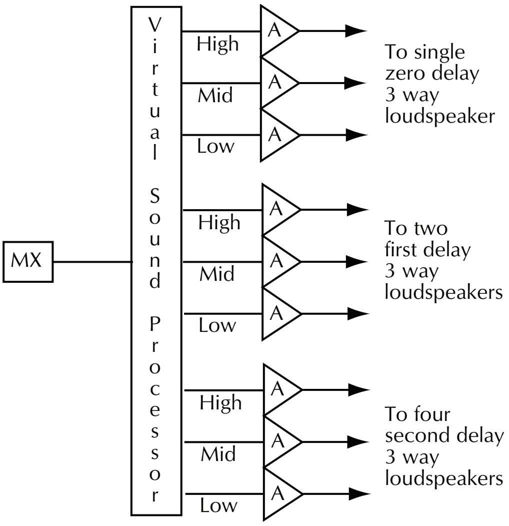

MX
Virtual Sound Processor
High
Mid
Low
A
To single zero delay 3 way loudspeaker
To two first delay 3 way loudspeakers
To four second delay 3 way loudspeakers

Figure 25-5. Virtual sound system processor treatment.

Following the trail blazed by Peavey, other manufacturers are now offering their versions of virtual sound processors.

Although not commonplace at the moment, in the not too distant future analog to digital conversion will occur at the system microphones with these signals being communicated directly to the system mixer or virtual system processor by optical cables making the system immune to both hum and radio frequency interference.

25.3 Device Interconnections

We will now explore the device interconnections required in the various treatments of the house of worship system discussed above. In the course of doing so we will identify the wiring techniques, connection techniques, and communication standards that will be the subject of detailed later discussion. Microphones typically require shielded twisted pair circuitry. In fixed installations this usually takes the form of a short flexible cable attached to the microphone by means of a female XLR-3 connector. This cable in turn is terminated by a male XLR-3 connector that is inserted into a wall or floor mounted receptacle fitted with a female XLR-3 connector. The microphone cable itself must be highly flexible and free of triboelectric effects. This is usually accomplished by employing a braided shield over an insulated twisted pair with the combination being covered by a woven fabric. This structure is in turn covered by a rubber or neoprene-insulating jacket. From the first receptacle, the microphone circuit continues within a conduit to a second receptacle mounted in the vicinity of the mixer location. The wiring within this conduit is usually in the form of an insulated twisted pair covered by aluminum foil and a drain wire. This foil-covered pair is in turn encased by an insulating jacket. The second receptacle is fitted with a male XLR-3 connector. A short cable similar to that attached to the microphone completes the circuit to the mixer input. Close attention must be directed to the grounding techniques employed in this wiring arrangement.

In the initial treatment of Fig. 25-2 and its variant of Fig. 25-3 all of the signals in connecting links are analog. All of the circuitry is balanced with the exception of that associated with the loudspeaker wiring. The loudspeaker wiring may or may not be balanced depending on the details of the amplifier output design. The output connector of the analog mixer is a male XLR-3 while the input connector to the equalizer unit may be a female XLR-3, screw terminals on a barrier strip, or "Euro" style terminal block inputs. The output of the equalizer can feature the same diversity of connectors with the XLR-3 being male rather than female. This array of connection possibilities is likely repeated on the signal delay unit. The power amplifiers will feature a similar choice of input connections while the output connection possibilities are binding posts, screw terminals on a barrier strip, or Neutrik loudspeaker connectors. The balanced circuit linking connections more than likely would employ a foil sheathed twisted pair of the same type employed in the microphone conduit. The primary consideration in loudspeaker wiring is the total wire resistance as compared with the impedance presented by the loudspeaker load. This is true for two reasons. Firstly, when the amplifier drives the loudspeaker directly, the circuit current, I, is common to both the wiring resistance, R_w, and the loudspeaker impedance, Z_l. This being the case, at any particular frequency it is possible to write the following average power relationships:

$$\begin{aligned} P_w &= I^2 R_w \\ P_l &= I^2 |Z_l| \cos\varphi \\ P_t &= P_w + P_l \end{aligned} \tag{25-1}$$

where,
P_t is the total average power,
φ is the loudspeaker impedance angle.

The real part of the loudspeaker impedance is given by

$$R_l = |Z_l| \cos\varphi \tag{25-2}$$

The fraction of the total average power that is delivered to the loudspeaker is then

$$\begin{aligned} \frac{P_l}{P_t} &= \frac{I^2 R_l}{I^2 R_w + I^2 R_l} \\ &= \frac{R_l}{R_w + R_l} \\ &= \frac{1}{\dfrac{R_w}{R_l} + 1} \end{aligned} \tag{25-3}$$

Eq. 25-3 clearly indicates that in order to deliver most of the power to the loudspeaker, the wiring resistance must be small as compared with the real part of the loudspeaker impedance. Long wire runs will thus require wire of large diameter in order to maintain the wiring resistance at a sufficiently small value. In many instances, particularly when dealing with large powers, rather than employing large diameter wire, it is more economical to employ a step-up transformer at the amplifier output and a step-down transformer at the loudspeaker location. In such instances, the wiring resistance is compared with a transformed real part of the loudspeaker impedance that is now n^2 times as large as before, n being the step down transformer's primary to secondary turns ratio. Common voltage values employed in such an application are 70.7 V, 100 V, and 200 V. Shielded cable is not required for loudspeaker wiring and is in fact undesirable. High voltage loudspeaker wiring, however, must be contained in conduit in order to meet electric code requirements. Secondly, loudspeakers are designed with the assumption that the driving source resistance is negligible as compared with the nominal loudspeaker impedance. Modern power amplifier source resistances certainly meet this requirement but the source resistance seen by the loudspeaker is that of the amplifier plus that of the loudspeaker wiring. A large value of source resistance as viewed by the loudspeaker will be detrimental to both the steady state response of the loudspeaker and more particularly to its transient response. For this reason also, direct connection between power amplifier and loudspeaker requires low resistance wiring. Higher resistance wiring can be tolerated in the high voltage system as the loudspeaker views a resistance equal to the wiring resistance divided by n^2.

The system treatments of Figs. 25-4 and 25-5 require digital signal interconnections. These connections require special cabling and communication protocols. The digital signals being communicated require a considerably larger bandwidth than that required by analog audio signals. One consequence of this increased bandwidth requirement is that lengthy interconnections must be treated as transmission lines. A transmission line, in order to operate properly, must have a well-defined characteristic or surge impedance and must be properly terminated in order to prevent reflections. The following pages are devoted to the grounding and shielding requirements of analog audio signal interconnections and detailed discussions of the communication standards, protocols, and cable structures required for digital audio signal communication.

25.4 Analog Interconnection Circuitry Types

There are basically two types of analog interconnection or link circuitry. Link circuitry is designated as either being balanced or unbalanced. Link circuitry in professional audio systems should in general be balanced with very few exceptions. One such exception is that of loudspeaker power circuitry which can be either balanced or unbalanced. Another exception occurs when provisions must be made to accept signals from unbalanced musical instrument sources as well as consumer or non-pro audio gear such as CD players, etc.

25.4.1 Balanced Circuits

A balanced circuit in its simplest form consists of two conductors that are symmetrical with respect to ground. Symmetry with respect to ground requires that the measured impedance between each conductor and ground must result in the same value of impedance. One of the consequences of this requirement is that the signal voltage measured between either conductor and ground will have an amplitude equal to one-half of the amplitude of the total signal voltage as measured between the conductors. Furthermore, the polarities of the two line to ground voltages are opposite. This situation is depicted in Fig. 25-6.

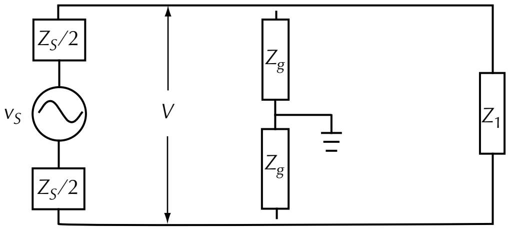


Figure 25-6. Basic balanced circuit.

In Fig. 25-6, v_s is a time dependent signal voltage generator having its own internal or source impedance and Z_l is the terminating load impedance for the balanced circuit. The virtual impedances Z_g represent the individual line to ground impedances measured with the signal voltage equal to zero and with the source impedance and terminating load impedance in place. The ground connection at the mid-point of the two Z_gs forces this point to be always at ground potential. The voltage between the upper conductor and ground will be $v/2$ while the voltage between the lower conductor and ground will be $-v/2$. That is, the two voltages are equal in magnitude but opposite in polarity.

One of the major attributes of a balanced circuit is its ability to reject common mode voltages. The word rejection as used here means that common mode voltages will not produce current in the load impedance Z_l. A common mode voltage is one where in the absence of the signal voltage v_s, both the upper conductor and the lower conductor will display identical voltages measured with respect to ground. Such voltages add to zero when summing around the entire circuit loop and thus do not produce current in the terminating impedance. There are occasions when a common mode voltage is purposely introduced into the circuit. One such occasion is that of supplying polarization voltage for capacitor microphones. The reader is referred to

Figs. 17-27 and 17-28 of Chapter 17 *Microphones*, Section 17.8 for circuit examples and discussion of this application.

25.4.2 Balanced Circuits and Susceptibility to EMI

Even when loudspeaker power linking circuits are excluded, analog audio linking circuits are called upon to handle a tremendous range of voltage levels. Microphone signals alone have normal operating voltage levels which vary from as low as –80 dBV up to about –20 dBV. So-called line level sources might exhibit normal values between –20 dbV to +4 dBV while mixer output levels can range up to +20 dBv or more. Electrical noise of any type appearing in microphone circuits is particularly troublesome because of the large degree of voltage amplification applied to such circuits.

Linking circuits that normally operate at higher levels must also be protected from electrical noise. There are quiet times and pauses in all program material. One can well be amazed how loud 100 mV of 60 Hz hum measured at a bass amplifier output can sound as played through an efficient loudspeaker system in a quiet auditorium. Even though balanced circuits inherently reject common mode signals whether purposely introduced or, under certain conditions, introduced as a result of electromagnetic interference (EMI), such circuits are not immune to EMI in general.

As an example, suppose the physical geometry of a linking circuit is similar to that of the drawing in Fig. 25-6. In this instance there are two long parallel conductors insulated from each other such that there exists a small but finite distance between the conductors, with the entire circuit forming an extended skinny closed loop. Further suppose that nearby there is an ac power raceway containing loosely separated power conductors with the raceway roughly running in the same direction as the link circuit. There will then be a predominantly 60 Hz alternating magnetic field in the vicinity of the link circuit. Any alternating magnetic flux that penetrates the area defined by the interior of the link circuit will induce an emf in the link circuit that obeys Lenz's law. This induced emf will produce a circulating current in the link circuit that attempts to oppose the externally applied flux change that brought it about. The unfortunate result is this induced voltage is not a common mode voltage but rather is of the same nature as is the desired signal voltage. The induced voltage will combine with the signal voltage with the combination of the two appearing across the load Z_l. The power line is also a nearby source of other noise components. The 60 Hz waveform is not a pure sinusoid and hence there will always be higher frequency harmonics present such as 120, 180, 240 Hz, etc. Silicon controlled rectifiers or similar devices as employed in lighting controls and dimmers introduce periodic damped radio frequency oscillations on the power line that can also induce noise signals in our example link circuit.

In modern times our audio equipment is constantly awash in a sea of electromagnetic radiation. In addition to the ordinary commercial AM, FM, and TV transmissions we now have cell phones, communication radios, wireless microphones, and a host of other electronic devices all of which are sources of electromagnetic waves that can lead to electromagnetic interference. This being the case, a few words about the nature of electromagnetic waves should be of value.

25.4.3 Brief Description of Electromagnetic Waves

A wave of any type is a physical disturbance that propagates with a characteristic velocity such that the disturbance is a function of both position and time. In the case of sound waves in air the disturbance is a variation in acoustic pressure and particle velocity occurring along the direction of propagation and hence is called a longitudinal wave. The acoustic pressure is called a scalar quantity as it has no direction. The particle velocity is a vector quantity having both magnitude and direction. The particle velocity oscillates back and forth along the direction of propagation of the wave.

Sound waves of course transport acoustic energy. Unlike sound waves in air, electromagnetic waves are transverse. The disturbance involves two vector quantities that are at right angles to each other in space with the pair in turn being at right angles to the direction of propagation of the disturbance. One of these vector quantities is called the electric field strength or **E** while the other is called the magnetic field strength or **B**. If the wave existed in air or in a vacuum and were to consist of a single frequency component, then both **E** and **B** would be undergoing sinusoidal oscillation in phase along their respective directions. A third vector quantity called the Poynting vector or **S** can be calculated from **E** and **B**. **S** always points in the direction of propagation and hence is always perpendicular to both **E** and **B**. The physical significance of **S** is that of the wave intensity or energy per unit area per unit time that is being transported by the wave. Electromagnetic waves can propagate freely in space and in air where the only loss is the normal spherical spreading with distance. The magnitude of the wave velocity in air is nearly the same as that in a vacuum for which the value is very close to 3×10^8 m/s.

Electromagnetic waves can also be guided by conductors such as on a transmission line or inside of hollow conducting pipes which are in fact called waveguides. The propagation velocities for guided waves are typically 60% to 70% of the free space value. In guided conditions there is no spherical spreading but attenuation does exist because of heat losses in the conductors. Electromagnetic waves can also propagate through other material media both insulating and conducting. The propagation velocity in insulators is typically 80% of the free space value and the waves are usually only weakly attenuated. The propagation velocity in the interior of good conductors such as copper is dramatically reduced as compared with free space conditions.

For example, at a frequency of 10^6 Hz an electromagnetic wave propagating through the interior of a slab of copper would have a propagation velocity of approximately 400 m/s. This is just a little more than the speed of sound in air! Furthermore, an electromagnetic wave propagating in the interior of a good conductor is rapidly attenuated as we will soon discover. The usual wave relationship between frequency, wavelength, and propagation velocity, ($\lambda f = c$), also applies to electromagnetic waves provided one employs the appropriate velocity for the guide conditions or medium in question.

Unlike sound waves in air, electromagnetic waves can be linearly polarized. The axis of polarization is that of the electric field. Fig. 25-7 illustrates the relationship between the field vectors for four different circumstances.

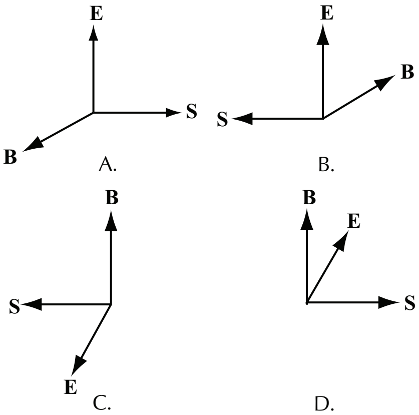


Figure 25-7. Cases illustrating the relationship between the field vectors for different polarizations and directions of propagation.

In each instance in Fig. 25-7 the field vectors form a mutually perpendicular set. In A and B the polarization is vertical with **E** oscillating along the vertical axis while **B** does so in the horizontal plane. In A the wave progresses to the right along the horizontal while in B it advances to the left along the horizontal. In C and D the polarization is horizontal with **E** oscillating along a horizontal axis while **B** does so along a vertical axis. In C the wave progresses to the left along the horizontal while in D the wave progresses to the right along the horizontal.

When an electromagnetic wave encounters a boundary surface or a change in medium a portion of the wave is reflected and a different portion is transmitted into the surface or into the new medium. The details of this process depend upon a number of factors. These factors are the angle of incidence on the surface, the polarization of the wave, the thickness of the second medium, and the electric and magnetic properties of the reflecting surface or medium. Of particular importance is whether the reflecting material is an electrical conductor or an electrical insulator. Highly conducting materials that are sufficiently thick reflect most of the incident electromagnetic energy particularly when the direction of propagation of the incident wave is perpendicular to the surface. The small portion of the wave not reflected by the conductor, the transmitted portion, is attenuated as it progresses deeper and deeper into the interior of the conductor. The attenuation of the field strengths, both electric and magnetic, as the transmitted wave penetrates the conducting material is given by

$$A = e^{-\frac{x}{\delta}} \tag{25-4}$$

where,

A is the dimensionless attenuation factor,
x is the penetration depth in m,
e is the base of the natural logarithm,
δ is the skin depth.

The skin depth, δ, is the penetration into the conducting material at which the attenuation factor becomes e^{-1} or 0.368. For all practical purposes the fields are completely attenuated after having traveled a distance of ten skin depths into the interior of the conductor. For a good conductor such as copper, the skin depth is given by

$$\delta = \sqrt{\frac{1}{\pi f \mu \sigma}} \tag{25-5}$$

where,

f is the frequency in Hz,

μ is the magnetic permeability in H/m,
σ is the electrical conductivity in siemen/m.

The frequency dependence of the skin depth when the conducting material is copper is exhibited in Fig. 25-8.

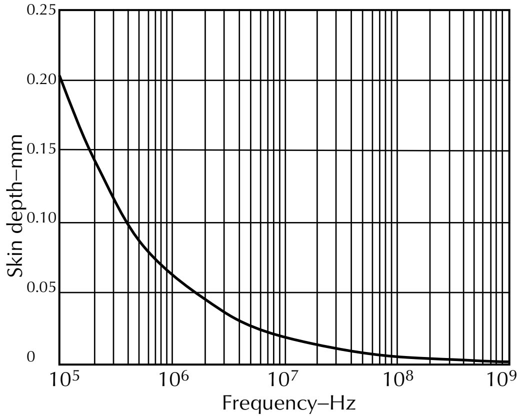


Figure 25-8. Skin depth as a function of frequency for copper.

Table 25-1 presents a tabulation of skin depth values for copper at a few selected frequencies.

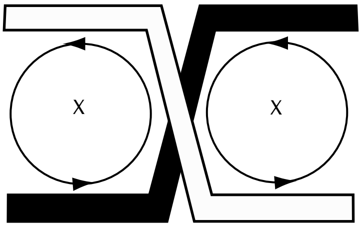
Table 25-1. Skin Depth Values

f in Hz	δ in mm
10^2	6.56
10^5	0.208
10^6	0.0656
10^7	0.0208
10^8	0.00656
10^9	0.00208

25.4.4 Shielding

Consider the task of completely shielding the circuit of Fig. 25-6 from the harmful influence of electromagnetic waves. This might be accomplished by encasing the entire circuit in a box made of copper or other good metallic conductor. If the walls of the box are as thick as 10 skin depths for the conducting material in question at all frequencies that are likely to be encountered, then two things will occur. Firstly, most of the electromagnetic energy incident from the exterior will be reflected at the outer surfaces of the box. Secondly, the weak transmitted residual wave energy will be completely absorbed as heat in the walls of the box before penetration of the total wall thickness occurs. The only wave energy in the interior of the box will be that of the signal that is being guided by the conductors constituting the balanced circuit. A perusal of Eq. 25-5, Fig. 25-8, and Table 25-1 indicates that at 60 Hz for a copper shield the wall thickness must exceed about ten times 8 mm, 80 mm, or about 3 inches. This would indeed be an expensive and impractical solution.

The objective, of course, is to make the balanced circuit insensitive to the influence of external fields that can generate loop as opposed to common mode voltages in the balanced circuit. Instead of forming the circuit from a pair of parallel conductors having a small but finite distance between them, the circuit should be formed of a tightly twisted pair of conductors as illustrated in Fig. 25-9.



Twisted Pair

Figure 25-9. Twisted pair exposed to a time changing magnetic field.

This figure, as well as the following discussion, is first presented in Chapter 17 *Microphones* and is repeated here for convenience. In Fig. 25-9 imagine that the twisted pair of conductors is replicated both to the right and to the left to form an extended circuit containing many twists with a small but uniform spacing between twists. Imagine also that in the vicinity a magnetic field is instantaneously directed into the figure as indicated by the Xs and that the field strength is increasing with time.

Examine the two closed paths as indicated by the circles. According to Lenz's law, the induced voltage acting in the loops has the sense indicated by the arrows. Now look at the white conductor in the upper left, the induced voltage in this portion of the conductor acts in the direction of the arrow adjacent to it. Compare that with the induced voltage in the white conductor in the lower right in which the induced voltage acts in the opposite direction.

The same analysis applied to the two similar segments of black conductor yields identical results. There is no voltage induced in the transposition region as the arrows in adjacent circles are oppositely directed. In practice, the magnetic field alternates but as it changes its direction of growth, the induction in the loops reverses direction also while the net voltage induced in the transposed conductors remains at zero. Complete cancellation of loop voltage will occur as long as there are many uniform twists in a distance equal to the wavelength of the offending field. Twisted pair audio cable of the type employed for microphones, etc., has about ten twists per foot. The frequency corresponding to a freespace wavelength of one foot is approximately 10^9 Hz. A twisted pair alone provides induced loop noise immunity from the very lowest frequencies up to about at least 10^8 Hz.

Twisted pair microphone cable is also supplied with a tightly braided copper shield. Even though such a shield does not provide total coverage, there are very tiny open spaces in the weave, the openings have dimensions that are small as compared with the wavelength of radiation up to about 10^{12} Hz. Radiation at lower frequencies than 10^{12} would view such a shield as being continuous. This braid is usually thick enough to provide total shielding above 10^7 Hz. Finally, if the braided shield is connected solidly to ground through the metal chassis housing the driving circuitry, the balanced conductors are shielded from capacitive coupling to external nearby power line or loudspeaker circuits. The shield in this instance provides a return path to the source for capacitively coupled noise currents.

25.4.5 Unbalanced Circuits

An unbalanced link circuit consists of an unbalanced line driving circuit and an unbalanced receiving circuit with the communication between the two circuits occurring over two conductors. The two conductors preferably are in the form of a coaxial cable that consists of an insulated central conductor and an outer braided shield. According to Maxwell's equations, coaxial cables themselves do not produce either electric or magnetic fields outside of the outer shield when the current in the inner conductor is equal in magnitude but oppositely directed to that in the outer conductor. Furthermore, when the shield supports an additional current there is no magnetic field in the region between the shield and the inner conductor attributable to the additional current. Unfortunately an unbalanced circuit is highly susceptible to electromagnetic interference at practically all frequencies and is particularly susceptible to power line induced interference. Consider the arrangement depicted in Fig. 25-10.

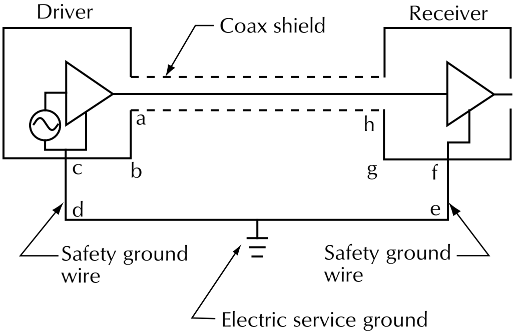


Figure 25-10. Unbalanced driver and receiver circuit.

In Fig. 25-10 an unbalanced line driver is connected to an unbalanced line receiver by means of a coaxial cable. Good grounding practice is employed in the line driver in that the signal circuitry ground connects to the power supply ground and the ac power circuit safety ground wire at only one point where all three are bonded to the metal shield or chassis enclosing the driver circuitry.

The same is true for the unbalanced line receiver. The driver output connects directly to the center conductor of the coax as does the receiver input. The driver-receiver signal circuit is completed through the outer shield of the coax that connects directly to the respective metal enclosures at each end. All of this is good practice and causes no problem in and of itself. This entire arrangement, however, is usually immersed in external alternating fields associated with both commercial ac power circuitry and radio frequency sources.

Consider the possible conducting path defined by the sequence of points a, b, c, d, e, f, g, h, and back to a. Any alternating field that produces a changing magnetic flux through the area defined by this path will induce an alternating current that will circulate around this conducting path. This induced current will exist in the shield of the coax. The shield of the coax may have just a small dc resistance but it also has an inductive reactance that increases linearly with frequency such that the total impedance associated with the shield itself can be appreciable. There then may be an appreciable induced noise voltage difference between the two ends of the shield equal to the product of the induced shield current with the shield impedance.

As a result the genuine signal voltage as measured between the center conductor and the shield at the driver end will differ from the received voltage measured between the center conductor and the shield at the receiver end. The received voltage will have been polluted by an amount equal to the noise voltage existing between the two ends of the shield. This phenomenon is called shield current induced noise or SCIN. The susceptibility to SCIN is an inherent shortcoming of unbalanced circuits.

25.4.6 Cables, Connectors, and the Pin 1 Problem

Fig. 25-11 displays two popular types of connector that are employed in analog audio linking circuits.

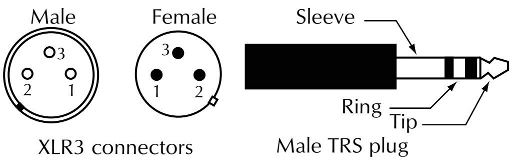


Figure 25-11. XLR-3 and TRS connectors.

A major attribute of the XLR type of cable mounted connector is the positive locking feature that is activated when the male and female are coupled together. Two conductor with shield cables are made with a male connector at one end and a female at the other so that standard length cables may be readily joined to produce any desired length. This is particularly helpful in non-permanent installations. In permanent installations cables are usually made up on site having the lengths required. Chassis mount versions of the XLR connector in both male and female also feature positive locking when mated with the appropriate cable connector.

The convention is that the connector associated with a signal source is male while that associated with a signal receiver is female. The figure also displays a male TRS connector for cable mounting. Cable mounted female TRS connectors are also available but are seldom used, as this type of connector does not have a positive locking feature. Short cables having TRS plugs at each end are often used with patch panels or jack fields. Cables having a female XLR at one end and a TRS plug at the other are often employed with devices that only have TRS chassis mounted input jacks. Cables and connectors intended for balanced linking circuits can be employed with unbalanced systems also whereas the converse is not true.

The standard pin assignment now in use has pin 1 connected to the cable shield, pin 2 connected to positive polarity signal, and pin three connected to negative polarity signal. When employed in unbalanced circuits, pin 1 is signal low or ground while pin 2 is signal high. The TRS or tip-ring-sleeve connector in the balanced application has the tip corresponding to pin 2, the ring corresponding to pin 3, and the sleeve corresponding to pin 1. When balanced XLR cables are used with microphones that require phantom power, the positive side of the dc phantom power circuit is fed to both pin 2 and pin 3 as a common mode voltage while the negative side of the phantom supply is connected to pin 1 as described in Chapter 17 *Microphones*, Section 17.5, and illustrated in Figs. 17-27 and 17-28.

25.4.7 The Pin 1 Problem

The pin 1 problem is of fairly recent origin and is principally the result of two factors. The first factor has to do with changes in manufacturing procedure. The second has to do with a loss of memory with regard to certain well-honed circuit engineering practices. In former times chassis mount XLR input and output connectors had metal housings and were mounted directly through mating holes of a heavy metal chassis that housed and shielded the associated electronic circuitry. All pin 1 terminals were either bussed through heavy gauge wire and then connected to the chassis or connected individually directly to the chassis. The reference ground for the active circuitry as well as the ground for the power supply had their own connections to the chassis. In permanent installations where custom cables were made up on site, pin 1 was connected to the cable shield only at the driving end for all subsequent balanced linking circuits. For non-permanent installations employing pre-made cables, pin 1 was connected at both sending and receiving ends of balanced linking circuits. Both types of installation were free of SCIN.

After the advent of printed circuit boards having automated parts insertion and soldering some, not all, manufacturers began to mount pcb XLR connectors on printed circuit board subassemblies and also began to connect pin 1 to the signal ground trace on the printed circuit board. Thus birth was given to the pin 1 problem as illustrated in Fig. 25-12.

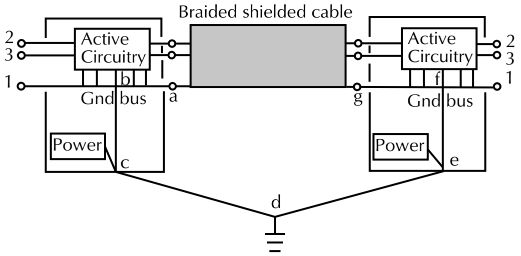


Figure 25-12. The origin of the pin 1 problem in balanced linking circuits.

The closed conducting loop in Fig. 25-12 described by the point sequence a, b, c, d, e, f, g, and back to a can have noise current induced in it by the same processes described with regard to Fig. 25-10. This same current exists in a portion of the printed circuit board signal ground trace or bus of both the send and receive circuits. The impedances of these traces are small but nevertheless significant and noise voltages will thus appear between various stages of the associated active circuitry. Good electronic engineering practice would not have allowed this to happen. The cure for this problem is illustrated in Fig. 25-13 that employs the time honored star grounding scheme.

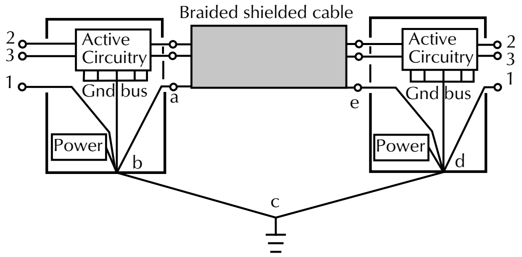


Figure 25-13. SCIN removed from signal ground reference by star grounding technique.

The SCIN loop of Fig. 25-13 now only involves the point sequence a, b, c, d, e, and back to a and there is no induced noise current on the active circuit ground bus.

Comments on Pin-1 Problems

Neil Muncy, a consultant in Canada, is a pioneer in the detection of Pin-1 problems, and has led the way to their correction in faulty designs. He wrote recently:

> *Regarding employing Pin-1 testing in the evaluation of new equipment, my long-standing policy is to categorically reject a device if it fails a Pin-1 test. Period. Send it back to the manufacturer along with a letter explaining why.*

Neil further comments:

> *It is up to the manufacturer to do whatever objective research is required to render their product subjectively free of Pin-1 issues. This is not a burden that should be dumped on the shoulders of innocent customers. Any manufacturer who cannot comprehend that when shield current applied to a Pin-1 produces an audible increase in output noise, should not be in the audio business!*

In today's marketplace many consumer-oriented devices end up being inserted into professional systems where everything from their connectors, impedances, and levels are not compatible with the other equipment.

25.4.8 Removal of SCIN from Unbalanced Circuits

Before the development of electronically balanced input and output circuits input and output transformers were universally used to accomplish the balancing tasks. Even today, such transformers are still employed to perform this task in certain very sensitive areas and particularly where ground isolation is required. A case in point is the removal of shield induced current noise in unbalanced circuits as depicted in Fig. 25-14.

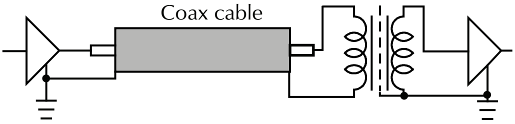


Figure 25-14. Transformer provided ground isolation.

Fig. 25-14 shows an unbalanced line driver connected to a coaxial cable at the sending end. The receiver end of the cable is connected to the primary winding of an input transformer that has a Faraday shield as a part of its structure. The secondary of the transformer is connected to the unbalanced receiver's input terminals and the Faraday shield is connected to the receivers signal ground reference terminal. Note that the cable shield is connected to ground at the sending end only so that there is no ground loop formed along which induced noise current might exist. The connectors at each end are usually RCA pin plugs and jacks. The employment of BNC male and female coaxial connectors would be preferred as they provide a positive locking feature.

25.4.9 Consumer Output to Balanced Input

On occasion it becomes necessary to connect a signal source having an unbalanced output to a mixer having balanced inputs. The signal source may be a CD player, radio tuner, non-pro tape player, or other similar devices whose output connector is usually an RCA pin jack. In such circumstances it is best to make up a braided shield twisted pair cable with an RCA pin plug on the sending end and an XLR on the receiving end with the wiring connections being done as given in Fig. 25-15.

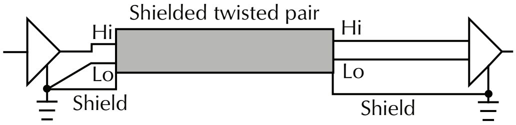


Figure 25-15. Adapter cable wiring for unbalanced to balanced.

25.5 Signal Cables—Analog Audio, Digital Audio, and Video

All electric signal transmission by means of a pair of conductors involves electromagnetic wave propagation with the conductors serving as boundaries for both the electric and magnetic fields involved. Knowledge of the electric field existing in the space between the conductors at some particular location along the conducting path and at some instant in time allows one to determine the potential difference between the conductors at that location and that instant in time. Similarly, knowledge of the magnetic fields at the surfaces of the conductors for a particular location and time allows one to determine the currents existing in the conductors at that location and instant of time. Maxwell's equations govern all electric, magnetic, and electromagnetic phenomena regardless of frequency and circuit dimensions.

In spite of, or perhaps because of, this universality, a great deal of mathematical expertise is required in applying Maxwell's equations to problems of a general nature. On the other hand, when the dimensions of the circuit are small as compared with the wavelength, Maxwell's equations reduce to the usual equations of lumped circuit element analysis involving Kirchhoff's laws I and II along with the usual definitions of capacitance, inductance, and resistance. Let the circuit in question be the connecting cable linking a single channel analog audio driver and receiver. For a high frequency limit of 20 kHz and assuming for the moment that the wave speed is the free space value of 3×10^8 m/s, the shortest wavelength would be 1.5×10^4 m or approximately 10 miles. Most such linking circuits in sound systems are orders of magnitude less than this so simple lumped circuit element analysis is valid and the cable can be modeled as displayed in Fig. 25-16.

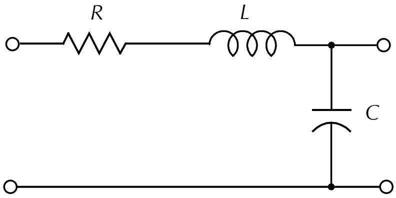


Figure 25-16. Lumped circuit model of analog audio connecting cable.

In Fig. 25-16, R is the total series resistance considering both conductors, L is the total series self-inductance considering both conductors, and C is the total capacitance between the conductors including the effect of a shield if present. This model neglects the effect of shunt conductance in parallel with the capacitance because with short cables this conductance is so small as to have a negligible effect. This model along with the source impedance of the driving circuit and the impedance presented by the load is sufficient for the calculation of attenuation as a function of frequency employing conventional circuit analysis techniques. One consequence of this model is that, aside from direction, the current in the upper conductor at a given instant is everywhere the same as that in the lower conductor. If the linking circuit dimensions were comparable to the wavelength, this would not be true and the model would be invalid.

25.5.1 Video and Digital Audio Signal Cables

The bandwidth for video signals extends from dc to at least 6 MHz. The upper frequency limit corresponds to a wavelength of approximately $(3 \times 10^8 \text{ m/s})/6 \times 10^6$ Hz or 50 m. This low value for wavelength is barely an order of magnitude greater than the length of many video linking circuits and in fact is comparable to or less than the lengths of many such circuits. Video linking circuits almost universally employ coaxial cables. A simple lumped circuit model cannot describe the behavior of the high frequency signals on such cables.

The problem is even more pronounced when dealing with digital audio linking circuits. Such circuits are recommended to meet standards as set forth in AES3. AES3 is a standard recommended by the Audio Engineering Society. The bandwidth specified by this standard for a single digital audio circuit communicating two channels of analog audio is a function of the individual channel sampling rate. The bandwidth specified by the standard for a sampling rate of 192×10^3 samples per second is 24.576 MHz. A maximum frequency of 24.576 MHz corresponds to a wavelength of approximately 12 m and thus all such linking circuits are comparable to or exceed the wavelength.

It is apparent that the description of the behavior of video and digital audio signals when propagating along connecting cables must be handled differently from that of analog audio signals in the typical sound system installation.

25.5.2 The Cable Problem

The description of the behavior of voltages and currents on extended cable structures is called the cable problem. An extended cable is one whose physical length is comparable to or larger than what we now know as the wavelength. The cable problem was first systematically studied and solved by Prof. William Thomson, later Lord Kelvin, in the middle of the nineteenth century. The problem was brought to Thomson's attention through the failure by others to establish telegraphic communication over submarine cables. The laying of successful telegraph cables on the ocean floor between Europe and the United States was the outstanding communication problem of that era. At the time, Thomson had an outstanding reputation in mathematical physics and occupied the chair in natural philosophy at Glasgow University in Scotland. Thomson's solution to the cable problem predated Maxwell's Treatise and probably contributed to Maxwell's thought in the ultimate formulation of his famous equations.

In formulating his model of the cable Thomson simply required that the structure of the cable be reasonably uniform. It mattered not whether the physical arrangement was that of just a pair of insulated wires uniformly spaced, a twisted pair in a shield, a coaxial cable, etc. The only constraint was that the series resistance, series self-inductance, shunt capacitance, and shunt conductance, all per unit of length, be uniform throughout the length of the structure. Thomson considered the series resistance and inductance as well as the shunt capacitance and conductance to be continuously distributed throughout the length of the system. He furthermore allowed for the possibility that the voltage between the conductors as well as the current in or between the conductors could vary significantly over small distances. Upon taking the axis of the system as the x-axis he was able to write equations relating the voltage, current, and the circuit parameters in the form

$$\begin{aligned}\frac{\partial v}{\partial x} &= -\left(Ri + L\frac{\partial i}{\partial t}\right)\\ \frac{\partial i}{\partial x} &= -\left(Gv + C\frac{\partial v}{\partial t}\right)\end{aligned} \qquad (25\text{-}6)$$

where,

v is the voltage between the conductors at any point x and time t,

i is the current in the conductors at any point x and time t,

R is the series resistance per unit length,

L is the series self-inductance per unit length,

C is the shunt capacitance per unit length,

G is the shunt conductance per unit length.

Eq. 25-6 is a pair of coupled equations in that both dependent variables or their partial derivatives appear in each equation. Uncoupling these equations leads to

$$LC\frac{\partial^2 v}{\partial t^2} + (RC + LG)\frac{\partial v}{\partial t} + RGv - \frac{\partial^2 v}{\partial x^2} = 0 \qquad (25\text{-}7)$$

and

$$LC\frac{\partial^2 i}{\partial t^2} + (RC + LG)\frac{\partial i}{\partial t} + RGi - \frac{\partial^2 i}{\partial x^2} = 0 \qquad (25\text{-}8)$$

Both the voltage as a function of position and time as well as the current as a function of position and time obey the same partial differential equation. If the circuit is excited at the sending end by a sinusoidal voltage source then the solution to the voltage equation will be of the form

$$v(x, t) = Ae^{-\alpha x}e^{j(\omega t - \beta x)} + Be^{+\alpha x}e^{j(\omega t + \beta x)} \qquad (25\text{-}9)$$

The solution for the voltage between the conductors is a function of both position and time and consists of two damped or attenuated waves. The term with the A coefficient describes a wave propagating from the source towards the receiver all the while being diminished in amplitude as x increases. Similarly the term with the B coefficient describes a wave propagating from the receiver back towards the source being attenuated as x decreases. This latter wave is a reflected wave. The coefficients A and B are in general complex and their values depend upon the conditions that exist at the input and the output ends of the cable. α and β are determined by the cable parameters and the operating frequency. Their structures in the general case are quite complicated. In the general case they are given by

$$\alpha = \sqrt{\frac{RG - LC\omega^2 + \sqrt{(R^2 + \omega^2 L^2)(G^2 + \omega^2 C^2)}}{2}} \qquad (25\text{-}10)$$

$$\beta = \sqrt{\frac{-(RG - LC\omega^2) + \sqrt{(R^2 + \omega^2 L^2)(G^2 + \omega^2 C^2)}}{2}} \qquad (25\text{-}11)$$

$$\alpha + j\beta = \sqrt{(R + j\omega L)(G + j\omega C)} \qquad (25\text{-}12)$$

25.5.3 Characteristic Impedance

If the cable were infinitely long there would only be a forward propagating wave. This would require the coefficient B to have the value of zero. In this circumstance, $v(x, t) = Ae^{-\alpha x}e^{j(\omega t - \beta x)}$. When this is substituted in the first of Eq. 25-6 and evaluated at the input end of the cable, it is learned that

$$i(t) = \frac{(\alpha + j\beta)}{(R + j\omega L)}v(t) \tag{25-13}$$

This leads directly to

$$\begin{aligned} \frac{v(t)}{i(t)} &= \frac{R + j\omega L}{\sqrt{(R + j\omega L)(G + j\omega C)}} \\ &= \sqrt{\frac{R + j\omega L}{G + j\omega C}} \\ &= Z_k \end{aligned} \tag{25-14}$$

Z_k is called the characteristic impedance of the cable and represents the input impedance of a cable of infinite length. If a cable of finite length is terminated in an impedance equal to Z_k the transmitted signal will be completely absorbed and there will be no reflection. Z_k is also called the surge impedance because even a cable of finite length that is improperly terminated will present this impedance to the driving source when the source is first connected to the cable. The source will continue to see an impedance equal to Z_k for a time interval T where T is the transit time required for a signal to propagate from the source to the receiver, be reflected, and then return back to the source. It is important to note that if the operating frequency is sufficiently high, ωL will be much larger than R and ωC will be much larger than G. In this circumstance, Z_k becomes purely resistive as given by

$$R_k = \sqrt{\frac{L}{C}} \tag{25-15}$$

The series self-inductance per unit length and the shunt capacitance per unit length depend principally upon geometrical factors. A larger spacing between conductors increases L and reduces C thus leading to higher values for the characteristic resistance.

25.5.4 Attenuation

α is known as the attenuation constant and in the general case is frequency dependent. The attenuation constant has the dimensions of reciprocal length. After traveling a distance equal to the reciprocal of α, a forward traveling wave will have its amplitude diminished by a factor of $e{-1}$ or 0.368. The voltage amplitude change in dB as a function of distance traveled can be written as

$$\begin{aligned} &20\,\mathrm{dB}\log(e^{-\alpha x}) = -8.686\alpha x\,\mathrm{dB} \\ &\text{or} \\ &-8.686\alpha\,\mathrm{dB/m} \end{aligned} \tag{25-16}$$

25.5.5 Phase Velocity

The phase angle of the forward traveling wave in the case of sinusoidal excitation at the source is $(\omega t - \beta x)$. The phase velocity of the signal on the cable is the velocity that an observer must have in order to observe a constant value for the phase angle. As time increases uniformly, the observer's x coordinate must also increase uniformly such that the phase angle retains a constant value. Therefore,

$$\begin{aligned} (\omega t - \beta x) &= \text{constant} \\ \frac{d}{dt}(\omega t - \beta x) &= 0 \\ \frac{dx}{dt} &= \frac{\omega}{\beta} \\ &= c \end{aligned} \tag{25-17}$$

The observer's velocity parallel to the cable axis is dx/dt and is called c. In order for all frequencies to propagate at the same velocity it is necessary that c be a constant independent of the frequency. Only under this condition will the transmission maintain a complex waveform's shape. This requires that β be directly proportional to ω. Only a cursory glance at Eq. 25-11 will indicate that this is not true in the general case. It can be made so, however. If the cable is constructed such that $GL = RC$ then Eqs. 25-10 and 25-11 simplify to

$$\alpha = \sqrt{RG} \tag{25-18}$$

$$\beta = \sqrt{LC}\omega \tag{25-19}$$

Eq. 25-19 then leads to a phase velocity that is independent of operating frequency as given by

$$\begin{aligned} c &= \frac{\omega}{\beta} \\ &= \frac{\omega}{\sqrt{LC}\omega} \\ &= \frac{1}{\sqrt{LC}} \end{aligned} \tag{25-20}$$

The operation of the cable under these conditions would be without distortion. Unfortunately, the requirement that $GL = RC$ is difficult to achieve in practice and an alternative approach is followed. Inherently, the shunt conductance is a small quantity and if $R/\omega L$ can be made small also except of course at very low frequencies, then Eqs. 25-10 and 25-11 become

$$\alpha = \frac{R}{2}\sqrt{\frac{C}{L}} + \frac{G}{2}\sqrt{\frac{L}{C}} \tag{25-21}$$

$$\beta = \sqrt{LC}\,\omega \tag{25-22}$$

In this case also the phase velocity is frequency independent except of course at the lowest frequencies. Typical phase velocities on cable structures range between about 50% to 80% of 3×10^8 m/s depending on conductor spacing and physical arrangement.

The conclusions above with regard to distortionless operation assume that the cable parameters are frequency independent. This can be approximately true over only a restricted bandwidth. When a cable is required to operate over a wide frequency range, account must be taken of the fact that the cable parameters are themselves frequency dependent. The basic Eqs. 25-6 through 25-17 still hold true but the attenuation constant, phase velocity, and characteristic impedance will all have values that vary with frequency. Of the four basic parameters, only the capacitance per unit length is essentially independent of frequency. The shunt conductance per unit length accounts also for the dielectric losses that are frequency dependent. The series resistance and self-inductance per unit length are frequency dependent because of a phenomenon known as the skin effect that will be discussed subsequently. As a result, wide-band signals transmitted over cables require equalization to compensate for frequency dependent attenuation and phase distortion. This equalization is usually applied at the receiving end for digital audio signals.

25.5.6 Skin Effect

Consider a long, straight homogeneous conductor of cylindrical cross-section that is far removed from other current carrying conductors. When such a conductor supports a direct current, the current is uniformly distributed over the entire cross-section of the conductor. A uniform direct current distribution, however, does not produce a constant static magnetic field within the interior of the conductor. Ampere's law requires that the magnetic field at the center of the conductor be zero.

Starting from the center where the magnetic field is zero, the magnetic field strength grows linearly with increasing radial distance from the center until it reaches its maximum value at the surface of the conductor where the radial distance equals the actual radius of the conductor. Outside of the conductor, the magnetic field strength begins to decrease being proportional to the inverse of the total radial distance measured from the conductor center. Fig. 25-17A illustrates the magnetic flux distribution in the interior of the conductor.

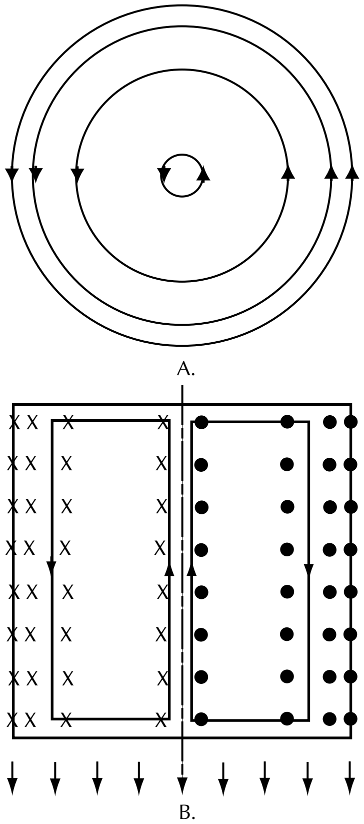


Figure 25-17. Magnetic field for direct current in a cylindrical conductor.

In Fig. 25-17A, the direct current is directed toward the observer. The concentric circles represent lines of magnetic flux. The spacing or density of the flux lines represents the magnetic field strength. Note that the lines are closer together as one nears the conductor surface indicating a stronger field in this region. In Fig. 25-17B, a slice has been taken through the center of the conductor. The slice extends for a small distance along the cylinder axis and the observer is looking down on this sectional view. The uniformly spaced arrows at the bottom of the figure represent the uniform distribution of direct current over the cross-section of the conductor. The arrows point in the direction of the current. The *x*s in the left half of the figure indicate that the magnetic flux is directed inward in this region whereas the dots in the right half represent magnetic flux directed outward in this region. Now imagine that the current is alternating at a low frequency and that instantaneously the current is increasing in the direction of the former direct current. This means that the magnetic flux is also increasing in the respective indicated directions. Consider now the lightly drawn rectangle in the interior of the left half of the figure. The magnetic flux through this rectangle is increasing inward. Lenz's law requires that the induced emf acting in this rectangle must oppose this increase in magnetic flux. It does so by driving a circulating current in the counter-clockwise direction as indicated by the arrows on the rectangle. In the right half of the figure the situation is just reversed because the magnetic flux is increasing outward. The induced current in the interior rectangle of the right half of the figure will be in the clockwise direction. Note that in each instance, the induced currents are oppositely directed to the original current in the inner regions of the conductor while being in the same direction as the original current in the outer regions of the conductor that are near to the conductor's surface. The result being that the net current is no longer uniformly distributed over the conductor cross-section. There is a higher current concentration near the conductor's surface at the expense of the inner regions. This phenomenon is known as the skin effect. The effect is more pronounced for conductors of large radius having a large electrical conductivity and a large magnetic permeability. For example, it is more pronounced in steel wire than in copper. The effect is more pronounced at high frequencies than low frequencies and becomes even more pronounced as the frequency continues to increase. Fig. 25-18 illustrates the current distribution in a 20 gauge copper wire with no other nearby conductors.

The relative current density for direct current has a value of unity at all points in the interior of the conductor as indicated by the horizontal line. This represents a uniform current distribution. Notice that at 10^5 Hz, the distribution has become slightly non-uniform with a higher current density near the conductor surface. The effect is even more pronounced at 10^6 Hz with practically no current near the center of the conductor. At even higher frequencies, the central region will become current free with all current concentrated in a shallow region or skin at the conductor's outer surface.

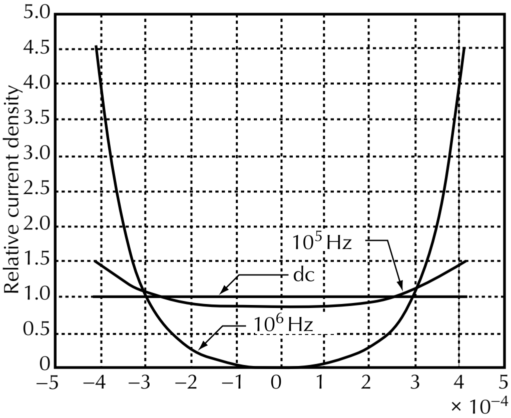


Figure 25-18. Skin effect in a 20 gauge copper wire.

Fig. 25-19 illustrates what happens to the conductor's resistance per unit length as the frequency is varied over a large range.

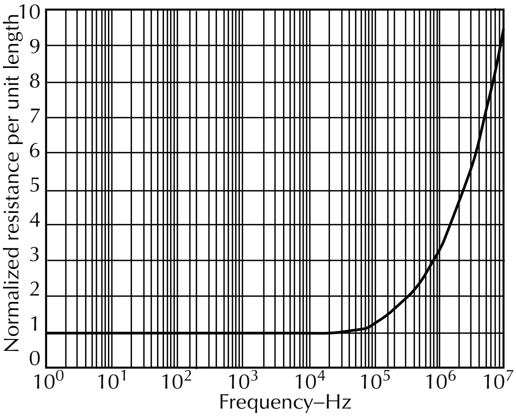


Figure 25-19. Ac resistance per unit length compared with dc value for 20 gauge copper wire.

Fig. 25-19 indicates that for an isolated 20 gauge copper wire the ac resistance is essentially the same as the dc value throughout the audio band with a significant departure beginning above 10^5 Hz. For larger gauge wires the departure will begin at lower frequencies.

When two conductors are close to each other and supporting equal but oppositely directed currents as in a cable, the proximity of one conductor to the other influences the current distribution in each conductor. In this situation, although the skin effect will be qualitatively the same in each conductor, the details will be modified dependent on the conductor spacing and other geometrical factors. The magnetic field between the conductors and at the conductor surfaces will also display frequency dependence such that the inductance per unit length of the conductors will depend upon both frequency and geometry. This behavior can be calculated from theory in closed mathematical form for certain circuit geometrical arrangements such as coaxial cables, balanced parallel conductors with or without shield, and others. For a difficult geometry such as a twisted pair in a braided shield, the frequency dependence of cable parameters is usually determined through direct measurement. For communication over a large bandwidth, the measurements must be carried out at numerous spaced frequency values in the communication band of interest.

25.5.7 Measurement of Cable Signal Propagation Properties

In calculating what changes a signal undergoes as it propagates along a length of cable, one needs to know the characteristic impedance of the cable, the source and termination impedances, the attenuation constant, and the phase constant. The inherent cable properties can be extracted from data provided by two relatively simple impedance measurements made at each frequency of interest. In this application, it is usual practice to take the origin of coordinates at the receiver end of a convenient length of cable. The source is then located at an *x* coordinate equal to –*l* where *l* is the physical length of the cable measured in meters. Employing Eqs. 25-6 and 25-9 it is relatively simple to determine the voltage and current at the source end of a length of cable under two different termination conditions. The chosen conditions are that of an open circuit condition at the receiver and a short circuit condition at the receiver. Upon taking the ratio of the voltage to the current at the source end one finds the impedance presented by the cable to the source under the given termination condition. For an open circuit cable this impedance is

$$Z_{open} = Z_k \frac{e^{+\Gamma l} + e^{-\Gamma l}}{e^{+\Gamma l} - e^{-\Gamma l}} \tag{25-23}$$

where,

$\Gamma = \alpha + j\beta$,

l = cable length.

On the other hand, for a cable that is short-circuited at the receiver the impedance measured at the source end becomes

$$Z_{short} = Z_k \frac{e^{+\Gamma l} - e^{-\Gamma l}}{e^{+\Gamma l} + e^{-\Gamma l}} \tag{25-24}$$

Multiplication of these two impedance values directly yields the square of the characteristic impedance so that

$$Z_k = \sqrt{(Z_{open})(Z_{short})} \tag{25-25}$$

The square root of the quotient of these two impedances after simplification yields

$$\sqrt{\frac{Z_{short}}{Z_{open}}} = \frac{1 - e^{-2\Gamma l}}{1 + e^{-2\Gamma l}} \tag{25-26}$$

Eq. 25-26 can be manipulated to obtain

$$e^{-2\Gamma l} = \frac{1 - \sqrt{\dfrac{Z_{short}}{Z_{open}}}}{1 + \sqrt{\dfrac{Z_{short}}{Z_{open}}}} \tag{25-27}$$

The expression on the right side of Eq. 25-27 is in general complex so it may be written in exponential form. The left hand side can also be rewritten to produce

$$\begin{aligned} e^{-2(\alpha + j\beta)l} &= e^{-2\alpha l} e^{-j2\beta l} \\ &= Me^{j\theta} \end{aligned} \tag{25-28}$$

where,

$\alpha + j\beta = \Gamma$,

M is the magnitude of Eq. 25-27,

θ is the angle of Eq. 25-27.

From Eq. 25-28 then

$$\begin{aligned} e^{-2\alpha l} &= M \\ -2\beta l &= \theta \end{aligned} \tag{25-29}$$

The attenuation constant can be determined directly from the first of Eqs. 25-29 by taking the natural logarithm of both sides. Care must be exercised in solving for the phase constant in the second of Eqs. 25-29. Dependent on the length of the cable sample used in the measurement relative to the wavelength of the signal on the cable, the angle of Eqs. 25-29 can have a value less than 2π plus or minus some integral multiple of 2π. For example, if it is known that the test cable sample is less than a wavelength long, then $0 < \beta < 2\pi$. Finally, having the value for the phase constant in hand, the last of Eqs. 25-17 can be solved for the phase velocity on the cable.

25.5.8 Sample Calculation

A one-meter sample of digital audio cable is operated at a frequency of 24.576 MHz in performing both an open circuit and a short circuit impedance measurement. The impedance observed at the source terminals when the receiver terminals are open is found to be $(7.5184 - j136.61)\,\Omega$. The impedance observed at the source terminals when the receiver terminals are shorted is found to be $(4.8597 + j88.303)\,\Omega$. When these values are substituted into Eq. 25-25 the characteristic impedance of the cable is found to be $110\,\Omega$. Next one takes the ratio of the short circuit impedance to the open circuit impedance and substitutes the result into Eq. 25-27. The ratio is found to be $(-0.64248 + j0.070933)$. Evaluation of the right side of Eq. 25-27 produces $(0.20384 - j0.92551)$. The magnitude of this complex quantity is 0.94770 while its angle is −1.3540 radians. These values are then substituted into Eqs. 25-29. Upon taking the natural logarithm on both sides of the first of Eqs. 25-29 and solving for α one obtains the value $0.02686\,\text{m}^{-1}$. This value can be used in connection with Eq. 25-16 to learn the attenuation rate on the cable is -0.2333 dB m^{-1}.

The second of Eqs. 25-29 can be solved directly for β to obtain $0.67701\,\text{m}^{-1}$. This value can be used in turn with Eqs. 25-17 to determine that the phase velocity on the cable at the operating frequency is $2.281 \times 10^8\,\text{ms}^{-1}$. This is 76% of the free space value. It is of interest to note that when measurements are made at one and two octaves below 24.576 MHz that the attenuation rate is $-0.1568\,\text{dB m}^{-1}$ and $-0.1007\,\text{dB m}^{-1}$, respectively illustrating the increasing attenuation rate with frequency brought about by the skin effect. These results are summarized in Table 25-2.

Table 25-2. Sample Calculation Summary

Quantity	Measured	Calculated
Frequency	24.576 MHz	
Z_{open}	$7.5184 - j136.61\,\Omega$	
Z_{short}	$4.8597 + j88.303\,\Omega$	
Z_k		$110\,\Omega$
α		$0.02686\,\text{m}^{-1}$
β		$0.67701\,\text{m}^{-1}$
Attenuation rate		$-0.2333\,\text{dB m}^{-1}$
c		$2.281 \times 10^8\,\text{ms}^{-1}$

25.5.9 To Terminate or Not to Terminate

The title of this section when posed as a question is a reasonable one. The current Audio Engineering Society standard that governs the electrical transmission of digital audio signals over connecting cables recommends a balanced system with both the source impedance and the receiver impedance being equal to the high frequency characteristic impedance of the connecting cable. This means that the cable "sees" its characteristic impedance at both ends. This is a good recommendation for the general case, particularly where the cable may not be a single continuous run but rather may be spliced.

Alternatively, the cable might be made up of two links joined by XLR connectors. Under either condition a reflection will occur at the discontinuity sending a return signal back towards the source. Such a reflection will be completely absorbed at the matched source and no subsequent reflections will occur. It should be mentioned, however, that a circumstance may arise where it might be desirable to terminate the cable in an impedance much higher than the characteristic impedance at the receiver end while maintaining a match at the source end. This circumstance is one where the cable is a long single continuous run. Such a long cable can introduce enough attenuation that the receiver will be voltage "starved." In this regard, it is worthwhile to compare the voltage at the receiver end under both matched and nearly open circuit conditions. Fig. 25-20 illustrates the situation as viewed by the source emf under the matched at both ends condition as well as the condition where the source is matched but the receiver has very high impedance.

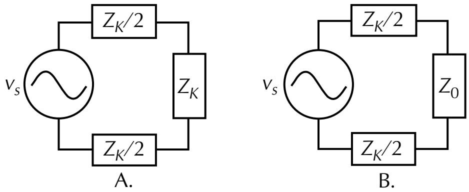


Figure 25-20. Equivalent circuit at source when the connecting cable is properly terminated, A, and terminated in an open circuit, B.

The equivalent circuits of Fig. 25-20 allow the rapid calculation of the voltage impressed across the input terminals of the connecting cable at the source location. In viewing Fig. 25-20A, remember that the input impedance of any length cable that is terminated in its characteristic impedance is the same as its characteristic impedance. As we are working with a balanced source, the source's internal impedance that is also equal to the characteristic impedance of the cable has been indicated in two parts with half in each leg of the generator. It should be clear from the figure that half of the open circuit voltage of the source appears across the input terminals of the cable. The general expression for the voltage anywhere on the cable is given by Eq. 25-9. In this instance, there is no reflection because we have a matched condition and B is thus zero. The sinusoidal open circuit voltage of the generator is given by

$$v_s = V_m e^{j\omega t} \tag{25-30}$$

Upon taking the origin of coordinates at the receiver, which places the source at $x = -l$ where l is the physical length of the cable, the voltage at the input end of the cable becomes

$$\begin{aligned} v(-l, t) &= A e^{\Gamma l} e^{j\omega t} \\ &= \frac{V_m}{2} e^{j\omega t} \end{aligned} \tag{25-31}$$

from which is learned that

$$A = \frac{V_m}{2} e^{-\Gamma l} \tag{25-32}$$

With the value of A in hand it is now possible to write an expression for the voltage at the receiver end where $x = 0$. Again employing Eq. 25-9,

$$v(0, t) = \frac{V_m}{2} e^{-\Gamma l} e^{j\omega t} \tag{25-33}$$

When $\Gamma = \alpha + j\beta$ is substituted into Eq. 25-33 the final result becomes

$$v(0, t) = \frac{V_m}{2} e^{-\alpha l} e^{j(\omega t - \beta l)} \tag{25-34}$$

The conclusion is that the voltage at the output is an attenuated and phase shifted version of the voltage at the input. Even in the absence of attenuation, the maximum voltage at the receiver is only one-half the voltage of the source. This result is to be compared with the one derived from the situation depicted in Fig. 25-20B. Here the receiver has high input impedance so the cable in effect is terminated in an open circuit at the receiver end. When there is an open circuit at the end of a cable, the current at this location will be zero and the incoming voltage signal is reflected in phase. This boundary condition makes the coefficients A and B equal. An inspection of Fig. 25-20B indicates that the voltage at the cable input terminals is given by

$$\begin{aligned} v(-l, t) &= V_m e^{j\omega t} \left| \frac{Z_o}{Z_k + Z_o} \right| \\ &= A(e^{\Gamma l} + e^{-\Gamma l}) e^{j\omega t} \end{aligned} \tag{25-35}$$

where,

$$\begin{aligned} Z_o &= Z_{open} \\ &= Z_k \frac{e^{\Gamma l} + e^{-\Gamma l}}{e^{\Gamma l} - e^{-\Gamma l}} \end{aligned}$$

After simplification Eq. 25-35 yields

$$A = \frac{V_m}{2} e^{-\Gamma l} \tag{25-36}$$

The voltage anywhere on the cable becomes

$$v(x, t) = \frac{V_m}{2} e^{-\Gamma l} (e^{-\Gamma x} + e^{\Gamma x}) e^{j\omega t} \tag{25-37}$$

At the receiver end of the cable $x = 0$, so the voltage there becomes

$$v(0, t) = V_m e^{-\alpha l} e^{j(\omega t - \beta l)} \tag{25-38}$$

Comparing Eq. 25-38 with Eq. 25-34 indicates that there is a factor of 2 or a 6 dB advantage in having a high impedance termination at the receiver. The conclusion is that for a very long continuous run of cable it may be advantageous to employ a receiver having an input impedance much higher than the characteristic impedance of the cable.

25.6 AES3

AES3 is an Audio Engineering Society recommended practice for digital audio engineering that covers the serial transmission format for two-channel linearly represented digital audio data. Linearly represented digital audio data can be interpreted as being pulse code modulation or PCM. Reference is made here to AES3-2003 as this is the latest revision as of this writing. Readers are urged to visit the AES web site in order to determine availability of later revisions.

25.6.1 Encoding Format

The encoding format is based on a block of 192 frames. Each frame consists of one sub-frame for audio channel one and one for audio channel two. The sub-frame encoding structure is based on 32 bits of which a maximum of 24 bits is devoted to audio sample data. The sub-frame can appear in two different forms, that for audio data of 20 bits or less and that for audio data above 20 bits but not greater than 24 bits. In either case, if the audio at hand has less bits than allowed then the least significant bits are padded with extra zeros to complete the record. For example a 16 bit audio sample would require an additional 4 zeros in the least significant bits positions. Fig. 25-21 illustrates this arrangement.

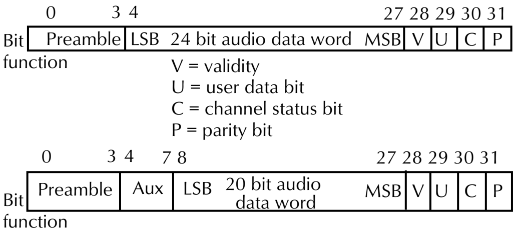


Figure 25-21. Sub-frame formats.

The preamble codes exist in three forms denoted as X, Y, and Z. The X preamble denotes audio channel one while Y denotes audio channel two. The appearance of preamble Z denotes the start of a new block consisting of 192 frames.

Channel status data is obtained by collecting the sequence of values contained in position 30, bit 31, of the respective sub-frames in each block of 192 frames. The appearance of preamble Z initiates this process. Thus for each channel one has a sequence of 192 bits that are formed into twenty-four bytes of 8-bits each. A great deal of information with regard to individual channel status can be conveyed by these twenty-four bytes and AES3 spells out in detail the significance of the various codes involved. Such information as channel identification, audio sample word length, audio sampling frequency, channel origin data, channel destination data, local sample address code, time of day sample address code, cyclic redundancy check character, and other similar information is all contained here. All data is transmitted serially from low order bits to high order bits.

25.6.2 General Transmission Characteristics

The bandwidth required for transmission is 128 times the frame rate and the frame rate is equal to the audio sample rate. At a maximum sample rate of 192×10^3 samples s^{-1} the required electrical bandwidth would then be 24.576 MHz. At a relaxed popular sampling rate of 48×10^3 samples s^{-1} the bandwidth requirement drops to 6.144 MHz. This has practical importance as quality video distribution amplifiers can handle this reduced bandwidth. AES3 does not require the use of transformers in electrical transmission but it does not preclude their use.

On the other hand the European Broadcast Union does require the use of transformers. AES3 allows for this by requiring that the data encoding technique employed has no direct current requirement when translated into an electrical signal. Furthermore it must be possible to recover the source clock at the receiver for synchronization purposes, and finally, the recovered data stream must be insensitive to the polarity of connections. This is facilitated by having all data beyond the preamble be encoded using the biphase mark technique. In this technique, each logic bit is represented by a symbol consisting of two consecutive binary states. The first state of the symbol must differ from the second state of the previous symbol and the second state of the symbol is the same as the first if the bit to be transmitted is logic 0. The second state of the symbol is different from the first if the bit is logic 1. This can be accomplished by employing a clock rate that is twice the bit rate. For example, suppose the logic code to be transmitted is the bit sequence 100110. The biphase mark encoding and the electrical signal representation of the encoding is illustrated in Fig. 25-22.

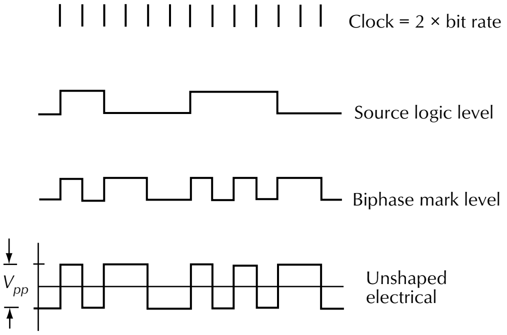


Figure 25-22. Channel coding scheme.

Preambles must have unique characteristics as they signal the start of a new sub-frame in the case of X or Y, or a new block of 192 frames in the case of Z. As such, the encoding for preambles must be structured so as to be readily distinguishable from the biphase mark technique employed for channel encoding. The preamble codes must also allow for clock recovery and electrically must be free of direct current. The first state of a preamble is always preceded by the second state of the parity symbol of the preceding sub-frame and the last state of a preamble is always succeeded by the first state of the LSB of the sub-frame with which it is associated. An example of preamble encoding as well as its associated unshaped electrical signal appears in Fig. 25-23.

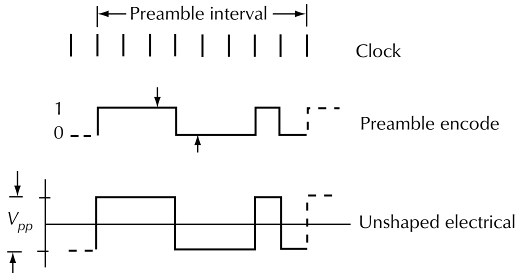


Figure 25-23. Preamble coding scheme.

In viewing Fig. 25-23 it should be remembered that the preamble occupies the first four of a sub-frame's allotted thirty-two bits and as such occupies eight clock intervals. The dotted line preceding the preamble represents the logic level of the second state of the preceding parity bit while the trailing dotted lines represent the transition to the first state of the successive LSB associated with the remainder of the sub-frame. The difference between the encoding of the preamble and the biphase mark technique is indicated by the two arrows on the middle drawing. The biphase mark encoding technique would require that transitions occur at the indicated positions. If one were to assign logic values to the individual clock intervals associated with the symbol in the middle drawing, the sequence would be 11100010. This coding would correspond to preamble X when the preceding parity-state happened to be a 0. If the preceding state had been 1, the preamble symbol would be different and correspond to the sequence 00011101.

25.6.3 Electrical Characteristics

The electrical characteristics as set forth in AES3 have been tailored to allow transmission of two channel digital audio over a balanced twisted pair shielded cable for a distance of 100 m without equalization for sampling rates up to 50 × 10^3 samples s–^{1}. Transmission over the same distance at higher sampling rates can be accomplished with appropriate equalization. The cable employed has a nominal characteristic impedance of 110 Ω for frequencies between 100 kHz and 128 times the maximum frame rate. The cable used in the previous cable sample calculation is a commercially manufactured cable specifically tailored for use under AES3. The cable driver is required to have a balanced source impedance of 110 Ω ±20% in the same frequency range. The cable receiver must also provide a balanced termination of 110 Ω ±20% in the same frequency range. The recommended connectors are XLR-3. The transmitted electrical pulses are bipolar but their rise and fall times between the 10% and 90% full amplitude points are to be limited to the range 5 ns to 30 ns. The peak to peak pulse voltage must fall in the range 2 V to 7 V. The measurement of these values is to occur across a 110 Ω resistor connected to the output terminals of the driver with no intervening cable. Rise and fall times and peak to peak voltage are illustrated in Fig. 25-24.

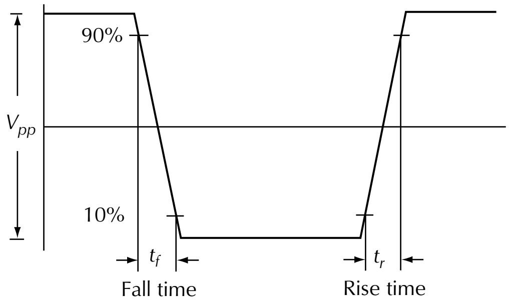


Figure 25-24. Rise time, fall time, and voltage peak to peak.

The process of generating, transmitting, and receiving serial digital audio is a synchronous process. Synchronization can be accomplished by having all elements of the system hardwired to a common master clock through separate clock reference circuitry. Alternatively, the clock may be located in the originating source with the signals so designed as to allow clock recovery at the receiver. This latter process is that covered by the AES3 standard. Accurate clock recovery is crucial to an error free processing of the received signals. Clock recovery hinges on the stability of the zero crossing positions of the received electrical waveform. Variation in the zero crossing timing is called jitter. The total jitter has two components, that associated with the internal clock and that associated with the intrinsic processing in the device. The standard sets forth limits for both intrinsic and clock jitter and details the methods to be employed in jitter measurements performed at the transmitter. Minimum requirements with regard to signal amplitude and timing necessary for successful recovery at the receiver are also set forth. The reader is referred to the latest version of AES3 for details as exact values are subject to revision as technology evolves.

25.6.4 AES Information Documents and Unbalanced Transmission of AES3

AES information documents are publications of the Audio Engineering Society in support of the AES standard documents. For example AES-2id is an information document for audio engineering that provides guidelines for the use of the AES3 interface. AES-3id-2001 is an AES information document for digital audio engineering covering the transmission of AES3 formatted data by means of unbalanced coaxial cables. Detailed information is provided with regard to cables, equalizers, adapters, and receiver circuits that allow the transmission of AES3 formatted data over distances up to 1000 m or in video installations that employ analog video distribution equipment. In the instance of a video installation employing analog video distribution equipment the frame rate is limited to 48,000 frames s^{-1} by the limited bandwidth of the analog video equipment. The coaxial cables employed have characteristic impedance of 75 Ω ±3 Ω above 100 kHz. The transmitter output signal requirements for transmission over unbalanced 75 Ω coaxial cable are summarized in Table 25-3.

Table 25-3. Unbalanced Output Signal Characteristics

Parameter	Minimum	Typical	Maximum	Unit
Output peak to peak	0.8	1.0	1.2	V
dc offset	—	—	< 50	mV
Rise time	30	37	44	ns
Fall time	30	37	44	ns

An AES3 transmitter with a balanced output can be converted to unbalanced operation through the use of an output transformer with one leg of the transformer secondary connected to the shield of the coaxial cable. In order for the transformer to provide the appropriate 75 Ω to 110 Ω impedance match it must have a primary to secondary turns ratio of 1.211 to 1. The AES3 voltage range at the primary of the transformer will be reduced by a factor of 1.211 as viewed at the transformer's secondary. This would exceed the maximum listed in Table 25-3, so further attenuation will be required. An attenuator that maintains a constant 75 Ω in both directions such as a T pad can furnish the required attenuation. An example of this conversion is shown in Fig. 25-25.

Suppose a balanced source conforming to AES3 has a peak to peak output of 5 V. Such a source can be comfortably converted for employment as an unbalanced 75 Ω source conforming to AES3-id by the circuit of Fig. 25-25. The voltage across the secondary of the depicted transformer when the circuit is terminated by 75 Ω will be the primary voltage divided by the turns ratio or 5 V/1.211 = 4.13 V. The T pad of the circuit when loaded by 75 Ω across the output terminals attenuates the 4.13 V by a factor of 4 thus producing an output voltage of 4.13 V/4 = 1.03 V. The operation of the original source will be completely normal as it will see a load of 110 Ω.

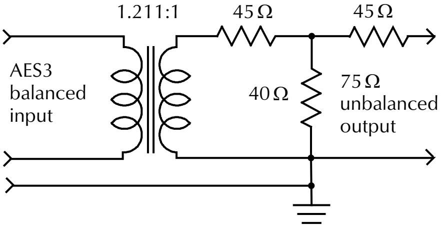


Figure 25-25. Balanced to unbalanced 110 Ω to 75 Ω with attenuation.

The connectors employed in electrical unbalanced coaxial cable operation are BNC connectors optimized for 75 Ω. The purpose of conversion to coaxial cable transmission other than to be compatible with analog video standards is to be able to transmit over large distances up to a kilometer. As such, cable loss becomes a significant factor. Cable losses expressed in units of dB/km are described by a relatively simple equation expressed as

$$A = a(f)^{1/2} + bf + g \tag{25-39}$$

where,
A is the loss in dB/km,
f is the frequency in MHz,
a, b, g are constants.

The loss characteristics for cables that are suitable for this application as extracted from AES3-id-2001 are listed in Table 25-4.

Table 25-4. Coaxial Cable Loss Characteristics

Cable	A 1 MHz	A 30 MHz	A 200 MHz	a	b	g
5C2V	8	47	126	8.5	0.031	–0.541
RG6A/U	7.9	48	135	8.4	0.082	–0.605
RG59B/U	14	62	175	9.6	0.177	4.24

25.6.5 Sony/Phillips Digital Interface Format

The Sony/Phillips digital interface format or S/PDIF conforms to AES3 with regard to data formatting and is employed in direct digital transfer from compact disc players and RDAT digital tape players. The electrical digital outputs of such devices are usually unbalanced and not intended for transmission over large distances. Small coaxial cables fitted with RCA pin plugs are employed in the link circuit. These devices also often feature optical interfaces with the same data formatting.

25.6.6 Microphones with Digital Outputs

Microphones having digital outputs are sometimes misleadingly referred to as being digital microphones. At the present writing at least such microphones have conventional analog microphone elements and analog preamplifiers combined with an ADC, an internal clock, and the associated circuitry necessary to present an electric serial digital output that conforms with AES3. Optical outputs as well will probably soon be in common use. This discussion will be restricted to microphones having electric serial digital outputs as a standard presently exists for microphones of this type. The current latest version of this standard as of the present writing is AES42-2001. One of the prime requirements for such microphones is the availability of a robust phantom supply capable of supplying currents much greater than those encountered when using analog output microphones. AES42 requires that the microphone receiver supply digital phantom power or DPP with a nominal dc voltage of +(10 + 0.5 – 0.1) V. Additional modulation of this voltage for other purposes to be discussed later involving +2 V ±0.2 V is allowed. The nominal available current required is 250 mA. A peak current of 300 mA is allowed when modulation is present. A circuit that allows the receiver to furnish phantom power to the microphone appears in Fig. 25-26A.

The cable involved in the circuit of Fig. 25-26A is twisted pair with shield as required by AES3 with the shield also acting as the phantom power return path. The transformer at the receiver location is center tapped to allow connection to the DP supply located in the receiver. The receiver might well be associated with an AES3 conforming input on a digital mixer. The transformer at the microphone location is center tapped on its secondary to provide the phantom power to the receiver. The phantom power appears as a common mode signal on the normal signal conductors and does not interfere with transmission or reception of the microphone signal data as long as the superimposed modulation on the phantom supply has a sufficiently small bit rate. This modulation is of the pulse width variety with a bit rate that is limited to the values listed in Table 25-5.

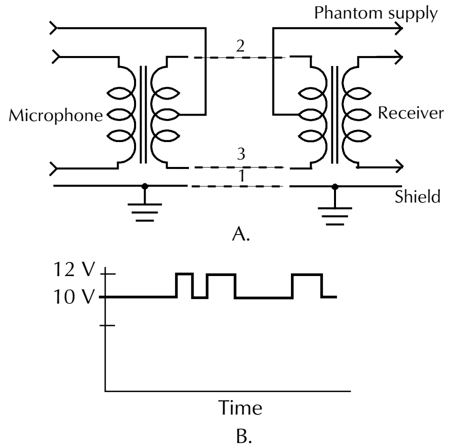


Figure 25-26. Phantom power circuit and modulated phantom power voltage.

Table 25-5. Modulation Bit Rate for Standard Sampling Rates

Sampling Rate—kHz	Bit Rate—bits s^{-1}
44.1	689.06
48	750
88.2	689.06
96	750
176.4	689.06
192	750

The modulation on the phantom power supply serves a variety of purposes depending on the mode of operation of the microphone. There are two possible modes of operation of the microphone as set forth in the standard. These modes are simply denoted as Mode-1 and Mode-2. In Mode-1 operation the clock internal to the microphone sets the timing and the receiver must recover the clock from the AES3 formatted data. In this mode the microphone may transmit information with regard to microphone ID and status of the operational parameters of the microphone through the employment of the user bits in the channel data stream.

The microphone control data is transmitted to the microphone from the receiver by means of the modulation that is imposed on the DPP. The instructions contained in this data can enable control of such factors as directivity pattern, low cut filter, microphone gain, signal limiter, and signal mute among others. In Mode-2 operation, the clock in the microphone is synchronized to the clock contained in the receiver. This synchronization is made possible through the issuance of frequent error correction commands contained in the modulation imposed on the DPP. In Mode-2 operation all of the features of Mode-1 with regard to microphone status and ID as well as control of microphone operational parameters are maintained. The use of Mode-2 operation is very important when employing several microphones in a given acoustical space along with a digital mixer. The digital mixer can furnish a master clock signal to all microphone inputs. In the absence of a common clock signal, the phases of the recovered audio signals from several microphones would be random and stereo imaging would be impossible. Mono operation could be seriously impacted as well when two performers using separate microphones are in close proximity. Mode-2 operation is absolutely necessary in order to maintain phase coherency between the recovered audio from several microphones.

25.6.7 Connectors for Digitally Interfaced Microphones

There are two schools of thought with regards to the connectors to be employed with digitally interfaced microphones. One school of thought would require complete compatibility with the XLR-3 connectors that are employed for analog interfacing. The other school would require complete incompatibility between the two applications. This latter position is based upon the differences between the phantom power supply standards for digital and analog.

These differences could well allow the possibility of damage to an analog interface if it is mistakenly connected to a digital interface receiver. At the present time there is no standard but AES42-2001 does make a recommendation that is a compromise position that can potentially satisfy both camps. The recommendation proposes an XLD connector that is based on an XLR-3 connector. The following is a quotation from informative Annex E of AES42-2001.

The XLD connector is based on the 3-contact XLR connector described in AES14 with the addition of grooves and user-insertable coding keys.

The chassis-mounting female connector includes a groove added to the inner contact insulator between contacts 2 and 3. In addition, a coding key may be added through the outer housing such that the end of the key protrudes into the space between the outer metal housing and the inner contact insulator. The key shall be positioned between contacts 1 and 2, close to contact 2, Fig. 25-27.

The male cable connector includes a groove that aligns with the optional coding key in the female connector. In addition, a coding key may be added through the outer shell of the male connector that aligns with the groove in the female connector.

To be fully coded, a connector shall have both a groove and a coding key. A half-coded connector has a groove only.

When XLD connectors are to mate only with other XLD connectors but never XLR connectors, then all XLD connectors shall be fully coded.

When complete interoperability of XLD connectors with XLR connectors is required, the coding keys shall not be inserted in either the male or the female XLD connectors.

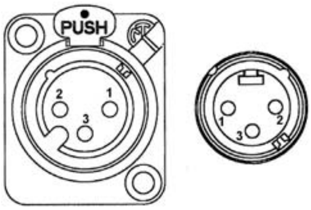


Figure 25-27. Fully coded XLD female chassis connector and male cable connector.

25.7 Computer Control and Communication of Digital Audio

Computers now permeate practically every field of human activity and audio systems are no exception. Penetration to the depth of individual sound system components is now a reality and, in the case of virtual sound processors, many former hardware elements are replaced by computer controlled DSP.

25.7.1 A Little History

In the late 1970s during the course of planning the sound systems for the new, sprawling Atlanta International Airport, F. B. Mewborn II, president of Baker Audio Associates, took a daring step. He realized that great economies could be achieved in system cost if all announcements could be made to flow through a switching central as is done in central telephone systems. In order for such a system to be successful it would be required that the switching be automatic and transparent to the user. It was realized that a computer could be made to perform the switching function provided that the originating address, the destination address, as well as the message were all in a form that the computer could understand. This would require analog to digital conversion of the audio source material prior to switching and digital to analog conversion after the switching had been accomplished with switching directional information being given by having the origin and destination addresses in digital form.

Prior to A-D conversion and after D-A conversion the audio source material could be handled in the conventional fashion employing microphones, amplifiers, equalizers, and loudspeakers then available. The author was privileged to serve as technical consultant for the development of this system.

While Baker Audio Associates was engineering this system it became apparent that the computer could also be made to perform a system diagnostic function through the insertion and observation of special audio test signals. The system went on line in 1980 with the diagnostic package being added shortly thereafter. At about this same time Innovative Electronic Designs began producing a similar system for application in airport sound systems and subsequently enlarged the scope of employment to any large-scale sound system. Up to this point all of the audio components other than the computer-operated switch were conventional.

In the ensuing few years Crown International Electronics, building on the digital electronics and programming experience it had acquired in the development of the TEF analyzer, began designing a new line of power amplifiers which would be amenable to both digital control and digital monitoring with the incorporation of a plug-in digital module. Parallel to the development of these amplifiers Crown also developed the computer interface and communication system necessary for interaction with these amplifiers. This work culminated in the Crown IQ System.

The Crown IQ System was structured on three levels. At the upper most level was the host computer and the IQ software. The host computer could be any IBM, IBM clone, or Macintosh computer that had a serial RS232, RS422, or RS4230 port. The computer acted as a monitoring and control station for the system. At the intermediate level was the Crown IQ interface that served as a communication device between the individual power amplifiers and the host computer. At the lowest level were the individual power amplifier plug-in microprocessor cards that were connected in a daisy chain by means of a single twisted pair to form a serial loop to the IQ interface. Communication between the interface and the individual amplifiers occurred at a baud rate of $38{,}400\,\mathrm{s}^{-1}$ so that the system operation appeared to occur almost in real time.

All of the normal manually controlled functions of each amplifier could be computer controlled by this system up to a total of 2000 two channel amplifiers. The outstanding feature of this approach, however, was that the actual operational status of each amplifier including on-off, input level, output level, distortion, and safe operating area was constantly monitored almost in real time.

The advent of the compact disc or CD player and RDAT tape recorders and similar devices led the Audio Engineering Society to establish standards for the transmission of two channel serial digital audio data. The first such standard appeared as AES3 published in 1985.

Finally, the modern era of computer control of sound systems began with the giant step taken by Hartley Peavey and Peavey Electronics Corporation in the introduction of the MediaMatrix® System that featured real time network communication and control of digital audio signals by Ethernet supported by CobraNet® hardware interface.

25.7.2 Networking of Digital Audio Data

The transmission and reception of serial audio data discussed thus far has been made by means of a synchronous process. In a synchronous process the frames of data are of constant size and data is transmitted and received at the same rate as governed by a single clock. The transit time over a particular connecting link is small and constant. This is in direct contrast to an asynchronous process. In an asynchronous process the size of a data block is not constant and a given data block is not sent at any particular time. The time of travel of a particular data block over a connecting link is not fixed and a single clock governing the overall process does not exist. Serial communication between a computer and a printer is an example of an asynchronous process.

Fortunately there exists a third type of process designated as being an isochronous process that will allow the transmission and reception of synchronous serial audio data over an interconnecting link where the link itself is operated asynchronously. In an isochronous process data is not transmitted at a fixed rate but there exists a fixed maximum transit time. Synchronous serial digital audio data can be successfully communicated via an isochronous connecting link through the inclusion of large data buffers at both the sending and receiving ends. The buffer at the sending end allows the retention of the proper sequence of serial audio data and the buffer at the receive end allows the re-synchronization of serial data at the appropriate rate.

The most commonly employed digital networking system for all purposes had its origin as far back as 1972 and is called the Ethernet. Ethernet has had continuous use and has undergone continuous improvements since its inception. This is an asynchronous system and as such can only be employed for the communication of synchronous serial digital audio data when further supported by isochronous adjuncts such as CobraNet®. For detailed information with regard to the history, operational characteristics, and terminology of Ethernet the reader is referred to the excellent article by Ray Rayburn referenced in the bibliography at the conclusion of this chapter.

25.7.3 CobraNet®

CobraNet® was invented by Kevin Gross and Rich Zweibel of Peak Audio. It has been licensed to so many major professional audio equipment manufacturers that it has become almost a de facto standard. CobraNet® is a combination of hardware (the actual physical interface), network protocol, and firmware. CobraNet® operates on switched Ethernet and in addition to the normal Ethernet services offers additional communication services. The additional services are isochronous audio data transport, sample clock distribution, and control and monitoring data transport. The interface performs all of the required isochronous to synchronous and synchronous to isochronous conversions along with all data formatting required for transporting real time digital audio data over an Ethernet network. In this regard, the Ethernet can be thought of as being a large conduit through which CobraNet® inserts and manages a group of sub-conduits as depicted in Fig. 25-28.

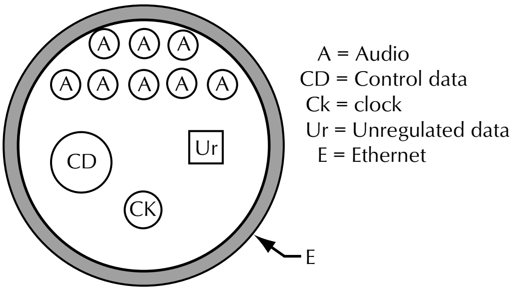


Figure 25-28. Ethernet with CobraNet®.

A typical CobraNet® interface offers the following features:

1. 100 Mbps full-duplex Ethernet interface.
2. Provides a backup Ethernet interface that can be connected to a redundant network for fault tolerance.
3. Quad synchronous output ports capable of supplying up to 32 total audio channels at 48 or 96 kHz sample rate with 24 bit resolution.
4. Quad serial input ports capable of receiving up to 32 audio channels at a 48 or 96 kHz sample rate with 24 bit resolution.
5. Clock source with jitter less than 1 ns.
6. High speed parallel host port interfacing to an optional external control processor.
7. Ethernet-based control, monitoring, and management.
8. Selectable low latency of 1.33, 2.66, or 5.33 ms across the network.
9. Asynchronous serial I/O port for bridge serial control data over Ethernet.
10. Status indicators for link, activity, fault, and CobraNet® conductor status for each Ethernet jack.

When two or more CobraNet® interfaces must communicate over the Ethernet only one interface supplies the common clock to be employed by all others. Electrical connection to Ethernet is via RJ-45 jack with integrated transformer isolation. Such a connection employs 2 of the 4 pairs in a Cat-5 cable. The serial audio I/O ports also feature RJ-45 jacks to accommodate RJ-45 plugs and Cat-5 cable. Cat-5 cable consists of 4 unshielded twisted pairs and is produced in two forms employing either stranded or solid wire. The nominal characteristic impedance of this cable is 100 Ω ±15%. The RJ-45 jack and plug are connectors that originated in the telephone industry as the terminology might suggest. They are compact modular connectors featuring eight contacts that in quality versions are gold-plated.

Bibliography

Audio Engineering Society Publications, Audio Engineering Society, Inc., 60 East 42nd Street, New York, New York 10165

AES3-2003

AES-3id-2001

AES42-2001

David J. Griffiths. *Introduction to Electrodynamics*, 2nd edition. Englewood Cliffs: Prentice Hall, 1989.

Ray Rayburn. “Digital Audio Interfacing and Networking,” *Handbook for Sound Engineers*, 3rd edition. Boston: Focal Press, 2002.

Appendix: Symbols and Abbreviations

a	absorption coefficient: a_S = Sabine a_N = Norris-Eyring a_R = room constant
$\bar{a}$	average absorption coefficient
A	ampere
ac	alternating current
AES	Audio Engineering Society
AIP	available input power
$\%AL_{CONS}$	percentage of articulation loss for consonants
AM	automatic mixer amplitude modulation
ANCA	ambient noise controlled amplifier
ANL	ambient noise level
ANSI	American National Standards Institute
APF	all-pass filter
ASA	Acoustical Society of America
BBF	band-boost filter
BPF	bandpass filter
BPST	bandpass sweep time
BRF	band-rejection filter
BW	bandwidth
c	velocity in many acoustic equations
C_L	coverage angle
CBF	constant bandwidth filter
CPB	constant percentage bandwidth
CPBF	constant percentage bandwidth filter
cps	cycles per second
dB	decibel(s)—a power ratio expressed on a logarithmic scale
dBA	A weighted sound-pressure level in decibels
dBm	decibels with a reference of 1 milliwatt (mW)
dc	direct current
dBV	decibels with a reference of 1 V
D_c	critical distance
DCD	directional control device
DFT	Discrete Fourier transform
D_I	directivity Index
D_m	measured distance
D_o	distance from talker to farthest listener
D_R	reflected sound distance
D_r	reference distance
D/R	direct-to-reverberant ratio
D_s	distance from talker to microphone
DVM	digital voltmeter
D_X	any given distance
D_1	distance between microphone and loudspeaker
D_2	distance from loudspeaker to farthest listener
D_{2SS}	D_2 distance for a single source
deg	degrees
DUT	device under test
ΔdB	level change in dB
ΔD_X	arbitrary level change associated with the distance D_X in the Hopkins-Stryker equation
E	voltage
E_{IN}	input voltage
EAD	equivalent acoustic distance
ECAN	electronically controlled ambient noise system
EFC	energy frequency curve
%Eff	percentage of loudspeaker efficiency
EIA	Electronics Industries Association
EIN	equivalent input noise
emf	electromotive force
E_o	open-circuit voltage
E_{OUT}	output voltage
EPR	electrical power required
ETC	Envelope Time Curve
Δ*EAD*	arbitrary level change associated with the distance *EAD* in the Hopkins-Stryker equation
F	distance from front of room to first listener
f	frequency
f_C	center frequency
FFT	Fast Fourier transform
FM	frequency meter frequency modulation
FO	frequency offset
FSM	feedback stability margin
FTC	frequency time curve
GLIT	ground loop impedance tester
GLR	graphic level recorder
GMT	Greenwich Mean Time
G_M	EIA microphone sensitivity rating
h	height of the loudspeaker
H	barometric pressure
HD	Haas distance
HPF	high-pass filter
h	hour
H_R	ceiling height at rear of room
HVAC	heat ventilating air conditioning
Hz	Hertz; cycles per second
I	current, intensity
I_a	acoustic intensity
IEEE	Institute of Electrical & Electronics Engineering
IR	Impulse Response
ITD	initial time delay
ISD	initial signal delay
$j, i, \sqrt{-1}$	an instruction to go to the 90° coordinate

k	kilo (1000)
K	Kelvin (temperature)
kg	kilogram
L_T	total sound level
L	level, length
L_{AMB}	ambient noise level
lb	pound
L_D	direct sound level
LEDE	Live End-Dead End
LF	low frequency
LFL	lower frequency limit
L_{in}	linear
log	logarithmic, base 10
ln	logarithmic, base e
$L_{P(EIA)}$	EIA loudspeaker sensitivity
LPF	low pass filter
L_X	loudspeaker sensi—not EIA (X may be 4 ft, 1 W 10 ft, 1 W 1 m, 1 W)
L_R	reverberant sound level
L_{RE}	level of early reflection
L_{SENSI}	loudspeaker sensitivity
LSI	large-scale integration
Ma	architectural D_C modifier
MAX	maximum
MDM	mix down monitors
Me	electroacoustic D_C modifier
MFP	mean free path
M	mega (1,000,000)
mic	microphone
min	minimum
MRT	modified rhyme test
MSC	miles of standard cable
ms	milliseconds
µs	microseconds
MTF	modulation transfer function
N	Newton
N	critical distance divisor (the ratio of acoustic power in the reverberant sound field to the acoustic power providing the direct sound field at the listener)
NAG	needed acoustic gain
ns	nanosecond
NBS	National Bureau of Standards
NBSA	narrow band spectrum analyzer
NC	noise criteria
NFI	null frequency interval
NFM	near field monitors
Np	Neper
NG	noise generator
NLA	noise level analysis
N/m^2	Newtons per square meter
NOALA	noise-operated automatic level adjuster
NOM	number of open microphones
NOMA	number of open microphone attenuators (an automatic mixer)
NPP	Nyquist phase plot
NQ_{min}	N times Q_{min}
NSCA	National Systems Contractor Association
OBA	octave-band analyzer
p	sound pressure
Pa	Pascal
PA	power amplifier
PAG	potential acoustic gain
PET	polar envelope time
Phon	loudness measure
P_D	phase delay
PF	power factor
PFC	phase frequency curve
Pink noise	equal noise energy per octave
PRD	primitive root diffusor
Program level	talker's level at the microphone without addition of meter lag
PRP	Pressure Recording Process
PZM	Pressure Zone Microphone
Q	directivity factor or electrical quality factor (used with resonance equation); see also R_θ
Q_{SS}	Q single source
Q_{REL}	Q Relative
Q_{avail}	available Q
Q_{min}	minimum Q for desired $\%AL_{CONS}$
QRD	quadratic residue diffusor
R	resistance or room constant (usually replaced today by $S\bar{a}$)
R_b	buildout resistor
R_{IN}	input resistance
R_M	match resistance
R_S	source resistance
R_L	load resistance
R_{OUT}	output resistance
RASTI	Rapid Speech Transmission Index
rad	radians
RC	resistor-capacitor
RFZ	reflection-free zone
RMR	center value of impedance range
rms	root mean square
RNG	random-noise generator
RPN	Reverse Polish notation (Lukasiewicz)
RPS	reflections per second
RTA	real-time analyzer
RT_{60}	reverberation time
R_θ	directivity ratio (also designated Q)
S	total boundary surface area
$S\bar{a}$	Sabins; usually supplied by material manufacturers in $S\bar{a}_s$
SA	signal alignment
Sabin	unit of absorption

SAG	sufficient acoustic gain
Sa_{obj}	Sabins per object
SBA	signal-biased amplification
SD	signal delay
s	seconds
sensi	sensitivity
SF	shelving filter
Sig Sync	Signal Synchronication
SLM	sound level meter
SNR, *S/N*	signal-to-noise ratio
S_p	microphone power sensitivity (with a reference of 10 dyn/cm²/1 mW)
SP	sound pressure
SR	sweep rate
sr	steradians, solid angle
STI	Speech Transmission Index
STP	standard temperature and pressure
S_V	microphone sensitivity with a reference of 1 dyn/cm²/1 V
T	time
TDS	time delay spectrometry
TEF	time energy frequency
TF	tracking filter
THD	total harmonic distortion
TIM	transient intermodulation distortion
TL	transmission loss
TN	thermal noise
TNL	thermal noise limit
TO	tracking oscillator
T_P	time period
TU	transmission unit
U	units
U_X	*X* number of units
UFL	upper frequency limit
µs	microseconds
V, v	volts
V	volume of a room
v	velocity
VA	volt amperes
VI	volume indicator
VOM	volt ohm meter
VU	meter calibrating unit
W	watt
W_a	acoustic watts
W_{act}	actual watts
W_{app}	apparent watts
W_e	electrical watts
White noise	equal noise energy per hertz
W_T	total power
X	unknown, Total reactance
X_C	capacitive reactance
X_L	inductive reactance
Z	impedance
$\lvert Z \rvert$	impedance (magnitude)
θ	phase angle
γ	Gamma

ρ	density of air
λ	wavelength, lambda
θ	horizonal coverage angle, theta
ϕ	vertial coverage angle, phi
Δ	Delta
Ω	ohms
π	pi
$<$	less than
$>$	greater than
≅	approximately equal to
≡	identical
≤	equal to or less than
≥	equal to or greater than
′	feet
″	inches

DIGITAL NOMENCLATURE

AAC	advanced audio coding
ADC	analog to digital converter
ASCII	American Standard Code for Information Interchanges
BAUD	a rate defined as Xbaud ´ Y bits ⁄ baud = Z bits/s. (Baud is changes of state/s) Named after French engineer Jean Maurice Emile Baudot
Binary	base 2, $\log_2$
Bit rate	bits ´ sampling frequency ´ channels
Bit	binary digit
BWF	broadcast wave format
Byte	8 bits
CBR	constant bit rate
CD	compact disc
Codec	compressor/decompressor
CPU	central processing unit
DAC	digital to analog converter
Denary	base 10, $\log_{10}$
Double word	32 bits
DSP	digital signal processor
DVD	digital versatile disc
ERB	equivalent rectangular bandwidth
FLOP	floating point operations per second
Fs	sampling frequency
GPU	graphics processing unit
HDTV	high-definition television
Hexadecimal	base 16, $\log_{16}$
HTML	hypertext markup language
IO	input – output
ISO	International Organization for Standardization
LSB	least significant bit
MDCT	modified discrete cosine transform
MIPS	million instructions per second
MIPW	million instructions per watt
MPEG	motion picture experts group

MSB	most significant bit
Nibble	4 bits
Octal	base 8, $\log_8$
PCM	pulse code modulation
PDF	portable document format
Quad word	64 bits
RISC	reduced instruction set computer
SNR	signal-to-noise ratio, $10\log_{10}(6 \times (2^{(b-2)})$
TNS	temporal noise shaping
$V/2^{\text{bits}}$	Quantization level
VBR	variable bit rate
WAV	wave format
Word	16 bits

ROOM MEASUREMENTS

cm	centimeter
D	dimension
Ft	feet or foot
In	inches
Kg	kilogram
Km	kilometer
L	length
L^2	area
L^3	volume
lb	pound
lbf	pound force
m	mile, meter, metric
mm	millimeter
yd	yard

Symbols
% efficiency 262
%AL_{CONS} 153, 254, 268, 269, 276, 277, 394
  choosing the correct equation 277
  effect of *SNR* 271
  factors affecting 272
  relationship to acoustic gain 278
  role of 277

Numerics
100Base-T Ethernet (Fast Ethernet) 544
10Base-T Ethernet network 543
20 log *x* chart 100
30/30 Guideline 437
3D surface model 459

A
Abbreviations 605
Abel, Niels Henrik 47
ABS 453
Absorption 177, 453
Absorption coefficient (*ABS*) 453
Absorptive materials 453
Ac voltage 114
Acoustic
  astigmatism 153, 470
  center 153, 470
  environment 171
  indoor environment 181
  origin 153, 470
  outdoor 179
  power 93
Acoustic astigmatism 153
Acoustic center 153
Acoustic dipole 310, 349
Acoustic distortion 311
Acoustic environment 267
  absorption 177
  ambient noise field 179
  atmospheric absorption 172
  diffraction 171
  diffuse (reverberant) fields 179
  Doppler Effect 175
  effect of space heater on flutter echo 177
  excess attenuation 171
  far free field 186
  far reverberant field 186
  flutter echo 177
  free fields 178
  indoors 181
  interfacing
    electrical output power 143
  inverse square law 172
  inverse-square-law level change 171
  Jesus addressed the multitudes 180
  mean free path (MFP) 184
  near field 186
  outdoor acoustics 179
  pressure fields 179
  reflection 171, 176
  refraction 171, 176
  reverberant sound field 184
  semireverberant fields 179
  sound absorption coefficients 182
  sound energy density 178
  sound fields 178
  sound velocity 173
  temperature dependant velocity 174
  typical wavelengths 175
Acoustic feedback
  analysis 556–558
  decay 557
  negative
    growth and decay 558
Acoustic gain 258
Acoustic gain parameters 258
Acoustic intensity 295, 308, 365
Acoustic kinetic energy 294
Acoustic measurements
  acoustic delay 197
  analytic signal 195
  comparison of an ear and a microphone 193
  constant percentage bandwidth analysis 203
  critical bandwidth 206
  decade exponents 208
  envelope time curve 195, 196
  equalization 192
  ETC plot 195
  fractional bandwidth filter analyzer 206
  frequency labels for audio components 208
  frequency resolution 198
  Hells Bells 201
  Heyser Integral Transform 193
  Heyser Spiral 196
  Hilbert Transform 196
  HVAC 203
  improper use of real time analysis 203
  impulse response 195
  impulse sources 192
  initial parameters 191
  label frequencies 208
  magnitude response 196
  measurement analyzers 193
  minimum phase response 196
  modulation domain analysis 193
  music 192
  NLA 203
  noise criteria 204
  noise criteria curves 203
  noise level analysis measurements 203
  non-minimum phase 196
  Nyquist
    plot 200
    trace of the complex signal 196
  phase and polarity 192
  phase response 196
  pinnea responses 207
  polar envelope time (PET) plots 201
  polarity 192
  processing TEF signals 194
  random noise 192
  signal synchronization 192
  site surveys 203
  speech 192
  steady-state sine wave 192
  swept sine wave 192
  TEF displays 196
  TEF instrument 193
  test sources 192
  three parameter measurement 198, 200
  time offset tracking filter 196
  time resolution 198
  trained human listener 191
  transducers 192
  undesired spurious energy 192
Acoustic origin 153
Acoustic potential energy 294
Acoustic power 485
Acoustic pressure 283, 286, 287, 290, 292, 294, 295, 297, 299, 304, 305, 307, 308, 310, 347, 351, 365, 371, 395, 408, 485, 498, 513
  far field 354
  near field 354
Acoustic pressure amplitude 293,

305, 310, 397, 405
Acoustic sound-pressure-level (L_p) 261
Acoustic Tests and Measurements 15
Acoustical analysis 579
Acoustics
  room 451
Active loudspeaker systems 428
Acustica 28
ADC 500, 538
Adiabatic constant 349
Adiabatic process 284, 286, 296
Admittance 124
AES3 541
  cable 599
  connector 599
  encoding 597
  transmission characteristics 598
AES3 vs. AoE 549
AES42 542
AES50 542
AES-EBU 541
AIP 316
Air
  atmospheric absorption 172
  density 173
  effect of altitude on the velocity of sound 175
Air density 347
Air particle 287, 292, 299, 309
Air particle displacement 291
Aliasing 39, 494
Alkmaion of Kroton 19
All Technical Services (Altec) 13
Altec 12
Alternating currents 113
Amoroso, Richard 19
Ampere, Andre Marie 9
Amplification
  voltage 88
Amplifier
  30/30 guidline 437
  apparent power 431
  as a voltage source 429
  burst testing results 435
  continuous sine wave power rating 434
  criteria for setting sensitivity 437
  digital signal processors 436
  equivalent amplifier size (*EAS*) 429, 430
  gain trading 439
  input sensitivity 436
  input sensitivity control 436
  load impedance 433
  low crest factor signals
    2 ohm loads 435
  mixer 142
  non-linear operation 428
  power from a voltage source 431
  power rating 428
  real world power generation 434
  setting sensitivity 436
  setting the amplifier's sensitivity 437
  voltage rating 429
Amplifier/loudspeaker interface 425
Amplitude 79
Amplitude response 382
Analog audio 550
Analog interconnections 583
Analog to digital conversion 497
Analog vs. digital audio 550
Analog waveform 535
Analog-to-digital converter (ADC) 535, 536
Analog-to-digital converters (ADCs) 35
Analysis
  constant percentage bandwidth 203
  feedback
    analysis 556–558
  pole and zero 508
Analyzer
  detecting reflections without analyzers 278
  fractional bandwidth filter 206
  real-time analyzer 559, 573
Angles 75
  coverage 153, 160
  epoch 75
  impedance 116
  improper coverage 275
  phasors 79
  pi 75
  polar lunes 156
  radians 67
  solid 67
  steradians 67, 156
Angular frequency 293, 300, 307, 311, 483, 486, 490, 492, 510
AoE 544
Apparent power 431
Arnold, H. D. 12
Articulation 270
Articulation index 271
Articulation losses of consonants in speech 268
ASCII
  code 46
  printable characters 46
Astrolabes 165
Asynchronous process 603
Atmospheric absorption 172
Atmospheric pressure 283, 287, 296
ATRAC (Adaptive TRansform Acoustic Coding) 38
Attack time 530
Attenuation
  excess 171
Audinate's Dante™ 547
Audio components
  frequency labels 208
Audio Engineering Society 35
Audio Perspective 12
Audio power 430
Audio-over-Ethernet (AoE) 543, 546
AutoCAD™ 460
Average loudspeaker sensitivity 438
Average power 118
Avogadro's number 283, 284, 287
Axial mode 241
Axial modes 240
Axon chips 19, 22

B

Baars, Bernard J. 20, 21
Babbage, Charles 12
Bais, F. Alexander 41
Balanced circuits 583
Ballou, Glen vii, 9
Bands 28
Bandwidth
  critical 206
  number of octaves 67
Bans 45, 48
Barks 28
Bartlett, Bruce 333
Bass, H. E. 171
Bauch, H. 28
Bauer, Ben 161
Bayes, Reverand Thomas 47
Bayesian probability theory 47
Bayesian theory 47

Beaty, H. Wayne 5
Beavers, Bob 161
Becker, Farrel 201, 274
Beethoven, Ludwig van 28
Bell Labs 12
Bell System Technical Journal 87
Bell Telephone Laboratories 87, 268
Bell, John (of Cern fame) 19
Benade, L. 374
Benson, J. E. 576
Beranek, Leo J. 240
Berger, Russ 248
Berkeley, Bishop 47
Berkeley, George 19
Bertram, Dr. Sidney 201
Bessel array 398
Bessel polynomials 510
Bickenstein-Hawking 35
Bilello, Charles 243
Binary 45
Binary choices 44
Binary codes 498
Binary numbers 45
Binaural recording 27
Biological memory 37
Biquad 526
Bissell, Chris 42
Bit rate reduction 39
Bits 45, 48
Black noise 40
Blind/deaf question 26
Bode plot 512, 513
Bodie, Hendrik 12
Bolt, Beranek & Newman 215
Bolt, Richard 217
Boner, Dr. C. P. 559
Bose-Einstein 22
Boundaries 309
Boyle, Robert 9
Bozak, Rudy 15
Brooks, Rollins 278
Brüel and Kjaer 14, 341
Burger, J. F. 268
Burst testing 435
Burst testing amplifiers 433
Bush, Vannevar 11
Butterworth polynomials 510
Byron, Lord 12

C

C 155
C_ 158, 159
Cable
  AES3 599
  analysis 591–597
  attenuation 592
  Cat-5 604
  miles of standard cable 87
  phase velocity 592
  propagation properties 595
  sample calculation 596
Callis, Maria 28
Campbell, George A. 11
Capacitive reactance 116
Capacitors 113, 125
Cat-5 cable 604
Category cable 541, 544
CATT-Acoustic™ 443
C_D 155
CEMI field theory 21
Characteristic impedance 592
Charles, Ray 26
Chart
  20 log x 100
  log 99
  polar response 154
Chebyshev polynomials 510
Circle
  sector 75
Circuits
  balanced 583
  dual 124
  matched 140
  open 140
  telephone 87
  unbalanced 587
Clarity-C 457
Clarity-C50 460
Clarke, Rick 469
Clausius, Rudolf 35
CLF Viewer 440
CobraNet™ 603
  interface features 604
Cobranet™ 547
Cochlear implant 26
Codec codes 538
Coding high quality digital audio 41
Cognitive computing 37
Comb filter 330
Comb filters 226
Components
  capacitors 113, 125
  inductors 113, 125
  resistors 113, 125
Compression 529
Computer aided system design
  3D surface model 459
  absorption 453
  absorption coefficient (*ABS*) 453
  absorptive materials 453
  AutoCAD™ 460
  CATT-Acoustic™ 443
  Clarity-C 457
  Clarity-C50 460
  counting balloon 450
  critical frequency char 454
  Dynamic Link Library (DLL) 447
  Early-Decay Time (*EDT*) 457
  early-reflected sound field 451, 455
  Echogram 460
  far field 444
  Finite Element Analysis (*FEA*) 460
  initial time gap (*ITG*) 455
  inverse square law 444
  late reflections or echoes 455
  late-reflected sound field 451
  line array of discrete sources 445
  log-squared *RIR* 452
  loudspeaker
    sensitivity 446
  loudspeaker arrays 447
  loudspeaker point source 444
  L_P 446
  modeling with absorption coefficients 453
  near field 444
  near field vs. far field 444
  objective of room modeling 451
  post-processing the *RIR* 452
  predicting room reflections 450
  radiation properties of a loudspeaker 445
  ray behavior 454
  ray tracing 449, 460
  realistic room models 457
  reverberant energy 456
  reverberant sound field 451
  RIR 456
  room impulse response (*RIR*) 451
  room model detail 449
  room modeling tips 460
  RoomCapture™ 457
  Sketchup™ 460
  sound behavior 454
  specular reflections 460

  spherical loudspeaker data 443
  T30 456
  wave behavior 454
  wire-frame model 458
Condensation 287, 289, 301
Conductance 124
Connectors
  AES3 599
  TRS 588
  XLD 602
  XLR 588
Conner, William K. 555
Conscious electromagnetic information field (CEMI) 21
Continuous sine wave power rating 434
Control code mnemonics 46
Controlling D_c 229
Conversion of Bits, Nats, and Bans 48
Cooper, Col. Jeff 20
Coulomb, Charles 10
Coverage angle 153
Craig, D. F, 576
Crick, Francis (of DNA fame) 19
Critical band 28
Critical bandwidth 30
Critical bandwidths 30
Critical distance (D_c) 157, 164, 228
Critical frequency (f_c) 215
Critical frequency chart 454
Crossover
  phase response 384
Crossover Network
  synthesized 390
Crossover network 378
Curves
  envelope time curve 568–569
    magnitude and phase 569
  house 561
  house curves 574

D

d'Alembert, Jean le Rond 11
D'Antonio, Peter 244, 248
D/R 253
D_0 254
D_1 158, 258
D_2 258
DAC 499, 538
Data rate 537, 539, 545
Data transport 537, 550
Davis, Art 15
Davis, Don 573, 576
Davis, Don and Carolyn 186, 279
dB 12
dBm 12, 91
dBV 88
D_c 258
D_c multipliers and dividers 229
Decade
  number in a frequency span 104
Decibel 71, 95
  20 log x chart 100
  acoustic intensity level (L_I) 93
  acoustic power 93
  adding decibel levels 97
  amplification 88
  apparent VU levels 104
  as a power quantity 90
  audio level 90
  bel 87
  calibrating a VI instrument 104
  combining 97
  combining levels of uncorrelated noise signals 98
  combining voltage 99
  dBm 91
  dBV 88
  deflection of the eardrum 105
  directivity factor 94
  distortion 106
  electrical power 90
  equivalent level (L_{EQ}) 96
  harmonic distortion 106
  inverse square law 93
  L_I 92
  log charts 99
  logarithm of a number to any base 100
  L_P 92
  L_W 92
  neper 87
  number of decades in a frequency span 104
  Ohm's Law 94
  older references 96
  percentages 109
  phon 105
  playback systems 108
  power 88
  power ratio 88, 89
  semitone intervals 101
  sensation unit SU 87
  subtracting decibels 98
  tempered scale 106
  the name for the transmission unit 87
  typical A-weighted sound levels 97
  VI meter 101
  voltage 88
  voltage amplification 88
  voltage amplitude 88
  voltage ratio 88
  VU impedance correction 102
  VU meter 101
  zero level 87
Decimal 45
Delay 467
Delta-sigma modulation 501
Designing for acoustic gain
  % efficiency 262
  %AL_{CONS} 254
  acoustic gain 258
  acoustic sound-pressure-level (L_p) 261
  D_0 254
  D_1 258
  D_2 258
  D_c 258
  direct-to-reverberant ratio (*D/R*) 253
  distance from the microphone (D_s) 254
  distance from the talker (D_0) 254
  EIA sensitivity 261
  equivalent acoustic distance (*EAD*) 253
  feedback stability margin (*FSM*) 255
  limiting parameters in sound reinforcement system design 260
  listener's maximum level ($L_{p(MAX)}$) 261
  loudspeaker efficiency 262
  loudspeaker sensitivity conversions 262
  measuring acoustic gain 259
  needed acoustic gain (*NAG*) 254
  nonlinear behavior 256
  nonlinear response 256
  number of open microphones (*NOM*) 255
  potential acoustic gain (*PAG*) 256, 259
  ratio 253
  regeneration 255
  required electrical power (*REP*) 261

  signal-to-noise ratio (*SNR*) 253
  sound radiation patterns 259
Device interconnections 582
D_I 222
Diagram
  pole-zero 509
Diamond, John 40
Difference equations 524–525
Differentiation 82
Diffraction 171
Diffusion 244, 245
Digital audio 536, 550
Digital audio cable 541
Digital audio formats
  100Base-T Ethernet (Fast Ethernet) 544
  10Base-T Ethernet network 543
  ADC 538
  AES3 541
  AES3 vs. AoE 549
  AES42 542
  AES50 542
  AES-EBU 541
  analog audio 550
  analog vs. digital audio 550
  analog waveform 535
  analog-to-digital converter (ADC) 535, 536
  Audinate's Dante™ 547
  Audio-over-Ethernet (AoE) 543, 546
  catagory cable 541, 544
  channel count 537
  Cobranet™ 547
  codec codes 538
  DAC 538
  data rate 537, 539, 545
  data transport 537, 550
  decodes 538
  digital audio 536, 550
  digital audio cable 541
  digital audio interface characteristics 541
  digital audio signal formats 539
  digital signal processing (DSP) 542
  dynamic range (DR) 535
  Ethernet 543
  Ethernet AVB 547
  Ethernet network 546
  Ethernet protocols 546
  eye pattern 536
  Finite Impulse Response filter (FIR) 543
  formatting 537
  Host ID 546
  Infinite Impulse Response (IIR) 542
  Internet Protocol (IP) 543, 547
  jitter 540
  jitter correction 540
  latency 537, 540
  local area computer network (LAN) 543
  MAC address 546
  Media Access Control (MAC) 543
  Metadata 537
  multi-mode fiber 545
  network bandwidth 543
  Network ID 546
  network node identifiers 546
  open standard 549
  Open Systems Interconnection (OSI) 547
  OSI Reference Model 548
  packet-switched network 543
  Quality of Service (QoS) 549
  quantization 535, 550
  reconstruction 536
  single-mode fiber 545
  Sony/Philips Digital Interface (S/PDIF) 540
  standard data formats 540
  synchronization 537, 539
  transport of digital audio data 537
  Virtual Local Area Network (VLAN) 548, 549
  word clock 539
Digital audio interface characteristics 541
Digital audio signal formats 539
Digital file size 38
Digital nomenclature 42, 607
Digital recording techniques 37
Digital signal processing (DSP) 542
Digital signal processors (DSP) 436
Digital systems 520–528
Digital theory
  aliasing 39
  analog-to-digital converters (ADCs) 35
  Bans 45, 48
  Bayesian theory 47
  binary numbers 45
  bit rate reduction 39
  Bits 45, 48
  black noise 40
  cognitive computing 37
  control code mnemonics 46
  digital file size 38
  digital nomenclature 42
  digital recording techniques 37
  digital versatile disc (DVD) 38
  dither 39
  dynamic range 36
  holographic universe 35
  latency 38
  Markov process 40
  Nat 37, 42
  Nats 45, 48
  over-sampling 39
  PCM (pulse code modulation) 35
  pink noise 40
  Planck constant 48
  Planck frequency 48
  Planck length 42, 48
  Planck mass 47
  Planck scales 47
  Planck system 47
  Planck time 48
  Planck units 47
  quantizing 35
  sampling 35
  Shannon entropy 41
  Shannon Space 36
  Shannon's theory 35
  Stochastic process 40
  text into binary, octal, hexadecimal 46
  thermodynamic entropy 41
  white noise 40
Digital versatile disc (DVD) 38
Digital voice recognition 22
Direct field (L_P) 449
Direct radiators 347
Direct sound field 451, 454
Direct sound level (L_D) 555
Direct sound pressure level (L_D) 222, 224
Directional control device (DCD) 470
Directivity
  far field 354
  function 355
Directivity factor 94
Directivity factor (Q) 153, 221
Directivity index D_I 157, 160

Directivity plots 155
  3-D plot 155
  frequency charts 155
  polar charts 155
Direct-to-reverberant ratio (*D/R*) 253
Disney 13
Displacement 80
Distance from the microphone (D_s) 254
Distance from the talker (D_0) 254
Distortion
  harmonic 106
  unnatural 572
Dither 39
Doak and Bolt delay 469
Dolby AC-3 38
Doolittle, Thomas B. 12
Doppler Effect 175
D_s 254
DSP 542
Dynamic Link Library (DLL) 447
Dynamic microphone equivalent circuit 323
Dynamic range 36
Dynamic range (DR) 535
Dynamics processing 528

E
Ear
  comparison to microphone 193
  deflection of the eardrum 105
Early-Decay Time (*EDT*) 457
Early-reflected sound field 451, 455
Echogram 460
Edelman, Gerald M. 20
Edison, Thomas 11
EIA noise spectra 426
EIA sensitivity 261
EIA sensitivity ratings 261
Eickemeyer, Rudolph 11
Eigen 239
Eigen mode 239
EIN 321
Einstein, Albert 19
Electric power 117
Electrical power required 260
Electrical Research Products, Inc. (ERPI) 13
Electromagnetic spectrum chart 5
Electromagnetic waves 584–586
  absorption 585
  polarization 585
  reflection 585
  skin depth 585
Embleton, T. F. W. 245
EMI 25
EMI susceptibility 584
Enigma code 47
Envelope time curve 269, 274
Envelope Time Curve (ETC) 14, 218
Environment
  acoustic 267
Equal loudness contours 30
Equalization 514
  global 517
  local 517
  system criteria 555
Equalizer
  passive 564
Equation
  dBm 91
  delay 467
  distance 467
  Sabine 215
  signal delay 466
  time 467
  velocity 467
Equation of state 283
Equivalent acoustic distance (*EAD*) 253
Equivalent amplifier size (*EAS*) 430
Equivalent rectangular bandwidth 29
Equivalent Rectangular Bandwidths (ERBs) 28
Ergodic processes 40
ETC 195, 269, 274, 568
Ethernet 543, 545, 604
Ethernet AVB 547
Ethernet network 546
Ethernet protocols 546
Euler, Leonhard 11
Euler's identity 405
Euler's theorem 77, 481
European International Advisory Committee 87
Evans, L. 171
Expansion 529
Eye pattern test 536
Eytomotic Research 27

F
Factors to watch for in rooms 231
Far field 444
Faraday, Michael 4, 9
Farmer, J. Doyne 41
Fay-Hall effect 468
Feedback
  acoustic
    analysis 556–558
    growth 556
  decay 557
  defined 556
  effects of surfaces 575
  feedback is a single frequency 570
  feedback stability margin 558
  negative acoustic 558
  oscillatory 556
  path 556
  positive 556, 558
  sources 569
  steady-state 572
  transient nature of acoustic feedback 556
Feedback stability margin (*FSM*) 255
Feynman, Richard 41
FFT 573
Field
  ambient noise 179
  diffuse 179
  far free 186
  far reverberant 186
  free 178
  multiplicity of 570–571
  near 186
  pressure 179
  reverberant 184
  semireverberant 179
  sound 178, 570
Field measurements 458
Fields 4
Filter
  active band-pass 564
  band-boost 562, 567
  bandpass 562, 567
  band-rejection 562, 567
  Bessel 386
  bridged-T 566
  broad-band
    combining-type band-rejection filter 565
  Butterworth 386

  Chebyshev 386
  criteria for band-rejection 562
  electric 379
  FIR 522
  high-pass 123
  Linkwitz-Riley 386
  low-pass 122
  minimum phase response 562
  minimum-phase 567
  narrow-band 564
  octave bandpass 121
  parameters 562
  recursive 524
  time offset tracking 196
  transfer characteristics 565
Filters
  IIR 524–526
  linear phase 527–528
Finite Element Analysis (*FEA*) 460
Finite Impulse Response filter (FIR) 543
Fink, Donald G 5
FIR 543
Fisher, Avery 15
Fleming, John Ambrose 4
Fletcher, Harvey 12, 27, 28, 87, 268, 279
Flutter echo
  effect of a space heater 177
Fogg Art Museum 216
Formatting 537
Fourier
  exponential series 477–482
  integral 483
  transform 484–485
  trigonometric series 477–481
Fourier transform 19
Fourier trigonometric series 478
Fourier, Jean Baptiste Joseph 477
Frequency 75, 175, 283, 292, 297, 302, 305, 307
  angular 75
  label 208
  resolution 198
Frequency dependent case 39
FSM 255
Fuller, Buckminster 19, 27

G

Gabor wavelets 19
Gabor, Dennis 13, 20, 40, 43
Gain structure 439
Galilei, Galileo 9
Gamma rays 5
Geddes, Earl 374
General Electric 11
General Radio 14
GenRad 341, 559
Geometric *Q* transformed to steradians (sr) 156
Geometrical acoustics 216
Gold Line TEF analyzer 468
Golden Ear 15
Goya, Francisco de 26
Green, Al 13
Gross Jr., Ervin E. 156
Gross, Kevin 604
Group delay 555
Guillmin, Ernst 5

H

Haas effect 243, 468
Hameroff, Stuart 20
Handbook for Sound Engineers 9
Harris, Cyril M. 171
Harrison, Ercel 13
Hartley, Ralph 40
Head related transfer functions (HRTFs) 27
Hearing versus listening 3
Heaviside, Oliver 4, 5, 11, 110, 502
Helmholtz, Hermann von 15
Henry effect 468
Henry, Joseph 10
Hertz, Heinrich 10
Hewlett Packard 559
Hewlett-Packard 14
Hexadecimal 45
Heyser spiral 13
Heyser, Richard 193
  Heyser Integral Transform 193
  Heyser Spiral 196
Heyser, Richard C. 9, 13, 25, 35, 39, 53, 561, 567
  TDS class 244
Hilbert Transform 196
Hilliard, John 15
Hippocratic School 19
Holograms 19
Holographic universe 35
Holography 49
Hooke, Robert 9
Hopkins-Stryker 157, 277
Hopkins-Stryker equation 259
Hopkins-Stryker—US and SI 222
Horn
  acoustic center 375
  acoustic origin 375
  astigmatism 374
  catenoidal 372
  compression drivers 376
  conical 370
  constant directivity 374
  cylindrical 371
  exponential 373
  polar measurements 375
  practical considerations 374
Horowitz, Vladimir 28
Host ID 546
How to Build Loudspeaker Enclosures 15
HP 35 14
Human brain 19, 36
Hunt, Frederick Vinton 9
HVAC 203

I

IBM 37
IEC noise spectra 426
Imaginary numbers 77
Imhotep 165
Impedance 115
  characteristic 592
  complex 137
  defined 135
  loudspeaker 129, 192
    electrical 364
  mechanical 350
  notation 136
  radiation 349
  Z 141
impedance angle 116
impedance bridge 126
Impedance diagram 116
Impedance of a real-world loudspeaker 433
Impedance tube 303
Improved reverberation time calculations 219
Impulse sampling 489
Indiana University 267
Inductive reactance 116
Inductors 113, 125
Infinite Impulse Response (IIR) 542
Information Theory 12
Information Theory Inference and

Learning Algorithms 36
Ingard 246
Initial signal delay gap (ISD) 240, 241
Initial time gap (*ITG*) 455
Input
  unbalanced to balanced 589
Instrument
  finding the velocity of sound with an analyzer 467
  high-quality wave analyzer 573
  improper use of real-time analysis 574
  measurement analyzers 193
  microphone calibrator 342
  real-time analyzer 576
  RTA application 561
  sweep oscillators 573
  TEF 193, 240
  TEF analysis 567, 576
  TEF analyzer 576
  TEF measurements 278
Integration 82
Intensity level 178
Interfacing
  available input power (*AIP*) 140
  complex impedance 137
  EIA microphone rating 139
  electrical gain of a system 142
  equalizer loss 144
  gain and loss blocks 141
  impedance
    defined 135
  impedance notation 136
  interfacing the electrical output power to the acoustic environment 143
  making reactance visible 136
  matched circuits 140
  mixer amplifier 142
  mixer output 140
  Nyquist
    plot 137
  open circuits 140
  technicians viewpoint 135
  Z (impedance) 141
  Z measurements 141
Interfacing systems
  ac current 113
  ac voltage 114
  admittance 124
  capacitors 113, 125
  circuits
    dual 124
  conductance 124
  filter
    high-pass 123
    low-pass 122
  impedance 115
    loudspeaker 129
  impedance angle 116
  impedance bridge 126
  impedance diagram 116
  inductors 113, 125
  Kirchhoff's Law 114
  network
    theorems 132
  networks
    constant resistance 128
    Zoebel 131
  octave bandpass filter 121
  phase angle 114
  power
    average 118
    electric 117
  power factor 119
  reactance
    capacitive 116
    inductive 116
  resistors 113, 125
  resonance
    parallel 124
    series 120
  root mean square 118
  susceptance 124
  two port devices 146
International Organization for Standardization (ISO) 38
Internet Protocol (IP) 543, 547
Intervals
  semitone 101
Inverse square law 172, 444
ISD 240, 241
Isochronous process 603
Isothermal process 284, 285, 296
It's from bits 38, 42
ITE (in-the-ear—at the eardrum) 27
ITE recordings 27
ITG 455

J

James B. Lansing 13
Jaynes, Edwin T. 47
Jeffreys, Sir Harold 47
Jet Propulsion Laboratory 13
Jewett, Frank 11
Jitter 540
Jitter correction 540
John, E. Roy 19
Jones, Doug 244
Joule, James Prescott 10
Journal of the Audio Engineering Society 14
Joyce, William 217

K

Kant, Immanuel 19
Keele, Don B. Jr. 354, 429
Keller, Helen 26
Kellogg, Edward W. 347
Kelvin temperature scale 322
Kelvin, Lord 9
Kendall, Gary 244, 245
Kennelly, Arthur Edwin 11
Kennelly-Heaviside layer 11
Killion, Dr. Mead 27, 193
Kircher, Athanasius 9
Kirchhoff, Gustav 5
Kirchhoff's Law 114
Klein, W. 279
Klepper, David 278
Klipsch and Associates 12
Klipsch, Paul 559
Kneser, Adolph 171
Knudsen, Martin 171
Koch, Christop 20
Kupfmuller, Karl 12, 40, 42

L

L Rad systems 40
Lagrange, Joseph Louis 11
Landauer, Thomas K. 41
Landauer's Principal 35
Lansing Manufacturing Company 13
Laplace transform 503, 520
Laplace, Pierre-Simon 11, 47
Large room acoustics
  comb filters 226
  controlling D_c 229
  critical distance (D_c) 228
  critical frequency (f_c) 215
  D_c multipliers and dividers 229
  D_I 222
  direct sound pressure level (L_D) 222, 224
  directivity factor (*Q*) 221
  Envelope Time Curve (ETC) 218

  ergodic specular enclosures 216
  factors to watch for in rooms 231
  geometrical acoustics 216
  improved reverberation time calculations 219
  large room frequency (F_L) 215
  levels in enclosed spaces 221
  Ma 230
  mean free path (*MFP*) 217
  N factor 230
  number of reflections (*N*) 217
  rate of decay of reverberant sound energy 218
  reflective path distance (*rp*) 226
  reverberant sound level (L_R) 224
  reverberation time (L_T) 224
  room absorption 230
  RT_{60} 226
  signal-to-noise ratio (*SNR*) 225
  sound intensity level (L_I) 220
  sound power level (L_W) 220
  sound pressure level (L_P) 220
  total power (L_W) 221
  what is a large room? 215
Large room frequency (F_L) 215
Late reflections or echoes 455
Latency 38, 537, 540
Late-reflected sound field 451
LEDE 240
  concepts 242
  design 248
Leibniz, Gottfried Wilhelm 22
Level
  acoustic intensity (L_I) 93
  adding 97
  apparent VU 104
  audio 90
  A-weighted sound level 97
  bel 87
  dBV 88
  decibel, the name for the transmission unit 87
  equivalent 96
  intensity 178
  inverse-square-law level change 171
  neper 87
  older references 96
  phon 105
  sensation unit SU 87
  sound pressure 178
  SU 87
  voltage amplitude 88
  zero 87
Levels in enclosed spaces 221
Limiting 530
Limiting parameters in sound reinforcement system design 260
Line array 400, 405
  near field 408
  processed 408
Line array of discrete sources 445
Linear phase 522
Listener
  trained 191
Listener's maximum level ($L_{p(MAX)}$) 261
Live End Dead End 240
Load impedance seen by amplifier 433
loading coils 11
Local area computer network (LAN) 543
Lochner, J.P.A. 268
Logarithm 60
  base e 76
  natural 76
  number to any base 100
Log-squared *RIR* 452
Lord Kelvin 191
Loss
  equalizer 144
Loudspeaker
  $\%AL_{CONS}$ 153
  acoustic dipole 349
  acoustic intensity 365
  acoustic pressure 347, 351, 365, 371, 395, 408
    far field 354
    near field 354
  acoustic pressure amplitude 397, 405
  adiabatic constant 349
  air density 347
  amplitude response 382
  array 392
    arraying techniques 395
    Bessel 398
    design process 393
    distributed 411
  average sensitivity 438
  bass vented enclosure 412
  Bessel array 398
  catenoidal horn 372
  circuit models 359–363
  compression driver 368, 376
  cone behavior 356
  conical horn 370
  constant directivity horn 374
  continuous line source 401
  coverage 153
  critical distance D_c 164
  crossover network 378
  cylindrical horn 371
  direct radiator spatial response 353
  direct radiators 347
  directivity
    far field 354
    function 355
  directivity factor 365
  directivity index D_I 157, 160
  distributed system 409
  dome radiators 358
  effect of combining unaligned loudspeakers in an array 472
  efficiency 363
  electrical impedance 364
  equal slope bandpass subwoofer 416
  equal-angle, weighted-area 161
  exponential horn 373
  frequency response of unaligned and synchronized horns 473
  half space radiation 353
  high density overhead 276
  horn 368
    acoustic origin 375
  horn acoustic center 375
  horn astigmatism 374
  horn polar measurements 375
  horn practical considerations 374
  hybrid arrays 411
  idealized geometry 160
  impedance 192
  large signal behavior 420
  length of conductor in gap 348
  line array 400, 405
    near field 408
    processed 408
  loudspeaker arrays 447
  magnetic induction in gap 348
  management systems 580
  maximun input voltage 426
  mechanical impedance 350
  mechanical limit testing 429
  minimum geometric *Q* 158
  minimum-phase 567
  misalignment 274
  missynchronization 470
  moving coil 348

- moving mass 348
- multi-loudspeaker arrays 570
- multi-way 438
- non-linear operation 428
- Nyquist diagram 380, 385
- overhead 409
- parameters 348
- pew back 276
- piston impedance function 349
- piston velocity 350
- point source 444
- polar response 471
- polarity 576
- power compression 427
- power response 384
- powered 439
- processed line arrays 405
- properly isolated 572
- properly shock mounted 572
- quality factor (Q) 351
- radiated power 358
- radiating surface area 347
- radiation impedance 349
- radiation properties 445
- radiation resistance 359
- radiator area 348
- resonance
  - free air 352
- resonant modes 358
- sensitivity 365, 446
- single source 393
- sound pressure level
  - axial 363
- split source 411
- static air pressure 349
- steerable line arrays 405
- surface acceleration 347
- surface displacement 347
- surface velocity 347
- suspension resistance 348
- suspension stiffness 348
- synchronization and alignment 469
- system 267
- transducers 192
- transfer function 379
- two misaligned 471
- under-balcony loudspeaker delay 469
- unequal slope bandpass subwoofer 418
- vented enclosure 420
- voice coil inductance 348
- voice coil resistance 348
- voltage rating 429
- volume velocity 347
- wire loss 582

Loudspeaker directivity
- $\%AL_{CONS}$ 153
- critical distance D_c 157
- directivity ratio 160
- equal-angle, weighted-area 161
- idealized loudspeaker geometry 160
- loudspeaker directivity factor (Q) 155
- minimum geometric Q by loudspeaker placement 158
- signal alignment 153

Loudspeaker efficiency 262

Loudspeaker management systems 580

Loudspeaker power ratings 425

Loudspeaker sensitivity conversions 262

Loudspeaker–room interactions 14

Lovelace, Lady 12

L_P 425, 430, 446

LSB 45

M

Ma 230, 277

MAC 546

MAC address 546

Mach, Ernst 19

MacKay, David J. C. 36

Maclaurin, Colin 47

Magiae Universalis 9

Magnetic phonograph cartridge 509

Magnitude vs. phase data 448

Marantz, Saul 15

Markov process 40

Martens, William 245

Martin, Bill 13

Martin, W. H. 87

Materials
- absorptive 453
- diaphragmatic absorbers 574
- quadratic residue diffusors 248
- Sonex 246, 278
- sound absorption coefficients 182

Mathcad 14

Mathematica 14

Mathematical operations
- addition 60
- antilogs 61
- division 60
- log multipliers 61
- logarithm 60
- multiplication 60
- powers 60
- roots 60
- subtraction 60

Mathematics
- accuracy 54
- addition 60
- amplitude 79
- anechoic chamber 71
- angles 75
- antilogs 61
- bandwidth 67
- base units 59
- complex numbers 64, 77
  - addition 78
  - products 78
  - quotients 78
- complex plane 77
- conversion factor 55
- decade 65
- decade calibration 65
- decibel 71
- differentiation 82
- displacement 80
- division 60
- energy 58
- epoch angle 75
- Euler’s theorem 77
- exponential notation 61
- factor label system 55
- force 58
- frequency 75
  - angular 75
- gains and losses 54
- imaginary number 77
- integration 82
- linear scales 66
- log multipliers 61
- logarithm 60
  - base e 76
- logarithmic scales 66
- math tables 72
- mathematical operations 60
- multiplication 60
- natural logarithm 76
- number of octaves 67
- numbers
  - complex 78
  - exponents 54

    - negative 54
    - ratios 54
    - reciprocals 54
    - roots 54
    - simple 54
  - octave spacing 66
  - percentages 70
  - phasor
    - addition 80
    - diagram 80
  - phasors 79
  - pi 75
  - power 58
  - powers 60
  - precision 54
  - pressure 58
  - *Q* 69
  - radian 75
  - radians 67
  - rate of change 80
  - ratios 70
  - real number 77
  - Renard Series 66
  - resolution 54
  - roots 60
  - sector 75
  - SI to US conversion 57
  - solid angles 67
  - spherical surface area 69
  - steradians 67
  - subtraction 60
  - trigonometry 75
  - US to SI conversion 57
  - velocity 80
  - work 58
- Matlab 14
- Maxfield's equation 270
- Maximum input voltage-MIV 425, 430, 437
- Maximum physical distance 253
- Maxwell, James Clerk 5, 9
- McFadden, Johnjoe 19, 21
- McIntosh, Frank 15
- Mean free path (*MFP*) 217
- Measurements
  - evaluating speech intelligibility 273
  - noise level analysis measurements 203
  - three parameter 198, 200
  - *Z* 141
- Measuring acoustic gain 259
- Mechanical impedance 300, 302, 305, 350
- Media Access Control (MAC) 543
- MediaMatrix® 581
- Mercedes Benz, Acoustics Lab 27
- Metadata 537
- Meter
  - apparent VU levels 104
  - calibrating 104
  - VI 101
  - VU 101
  - VU impedance correction 102
- Mewborn II, F. B. 602
- MFP 184
- MGM 13
- Microphone
  - *AIP* 316
  - axis 330
  - boundary 330
  - cables 339
  - calibrator 342
  - capacitor 324
  - carbon 324
  - cardioid 330
  - comb filter 330
  - comparison to ear 193
  - compression and expansion 338
  - connector
    - digital 602
  - connectors 339
  - digital output 601
  - directional characteristics 328
  - dynamic 322
  - EIA rating 139
  - *EIN* 321
  - measurement 341
  - moving coil 324
  - noise
    - thermal 317
  - noise figure 320
  - noise sources 319
  - Nyquist noise 317
  - passive filter set 316
  - phantom power 339, 340
  - piezoelectric 325
  - polar patterns 328
  - polarity 575
  - power
    - noise 318
  - pre-emphasis and de-emphasis curves 337
  - pressure 325
  - pressure gradiant 325
  - pressure gradient 325, 326
  - pressure zone 333
  - proximity effect 327
  - PZM 333
  - ribbon 325, 330
  - sensitivity 315
  - *SNR* 320
  - source impedance 323
  - speech intelligibility 267
  - wireless 335
- Millikan, Robert 11
- Milton, John 26
- Minimum phase 513
- *MIV* and *EAS*
  - combining 439
- *MIV* test setup for internally powered loudspeaker 428
- *MIV-burst voltage* 429
- MIV_{peak} 429
- MIV_{rms} 429
- *MIV-thermal* 429
- Mixer
  - output 140
- Modal decay rates 239
- Mode
  - axial 240, 241
  - damped and undamped 239
  - Eigen 239
  - oblique 240, 241
  - small room modes 239
  - tangential 240, 241
- Modeling with absorption coefficients 453
- Modes
  - proximity 575
- Modular transfer function 268
- Modulation
  - delta sigma 501
- Moir, James 246
- Moravec's warning 41
- Morrison, Ralph 5
- Morse 246
- Morse, Philip 217
- Motion Picture Expert Group (MPEG) 38
- Motivations 25
- Mozart, Wolfgang Amadeus 28
- MSB 45
- MTF 268
- Multi-mode fiber 545
- Multi-way loudspeaker 438

N

- *N* factor 230
- *NAG* 254
- NASA 13

Nat 37, 42
Nats 45, 48
Near field 444
Near field vs. far field 444
Needed acoustic gain (*NAG*) 254
Neper 87
Network bandwidth 543
Network ID 546
Network node identifiers 546
Networks
  constant resistance 128
  crossover 378
  theorems 132
  Zoebel 131
Neumann, John von 41
Neuron chips 19, 22
Newton, Isaac 10
Newton, Sir Isaac 47
NLA 203
Noble, Jim 15
Noise
  criteria 204
  pink 561
  random 192
  white 561
Noise control 555
Noise Criteria curves (NC) 560
Noise figure 320
Noise Measurement Handbook 160
Noise power 318
Noise signals
  combining levels 98
Noise sources 319
NOM 255
Nonlinear behavior 256
Non-linear dynamics processing 529
Nonlinear response 256
Non-minimum phase 513
Non-planar wave motion 297
Non-statistical spaces 237
Null frequency interval (*NFI*) 470
Number base 498
Number of open microphones (*NOM*) 255
Number of reflections (*N*) 217
Numbers
  complex 64, 77, 78
    addition 78
    products 78
    quotient 78
  imaginary 77
  real 77
Nyquist
  plot 137
Nyquist diagram 380, 385
Nyquist display 13
Nyquist noise 317
Nyquist, Harry 12, 37, 40, 568

O

Objective of room modeling 451
Oblique mode 240, 241
Octal 45
Octave
  number of 67
  spacing 66
Octave bandpass filter 121
Ohm’s Law 94
Olson, Harry 14
One Kelvin 35
Open standard 549
Open Systems Interconnection (OSI) 547
Origins in Acoustics 9
Orrery 165
Orwell, George 22
Osborne, Tom 14
OSI Reference Model 548
Over-sampling 39

P

Packet-switched network 543
PAG 256, 259
Parseval’s theorem 485
Particle acceleration 290
Particle displacement 287, 288, 290, 292, 299, 302
Particle velocity 287, 288, 290, 292, 297, 306, 308
PASC (Precision Adaptive Sub-band Coding) 38
Passive filter set 316
PCM (pulse code modulation) 35
Peavey, Hartley 603
Peerless Manufacturing 13
Penrose, Roger 20
Percentages 109
Performer 267
PET 201
Peterson, Arnold P. G. 156
Peutz, V.M.A. 268, 269, 278, 279
Peutz, Victor 26
Pharaoh Zoser 165
Phase 192
  linear 522
  minimum 513
  non-minimum 196, 513
  velocity 592
Phase angle 114
Phase response 384
Phase velocity 297, 302, 311
Phasor
  addition 80
  diagram 80
Phon 105
Phon level 29
Phons 28
Phonurgia Nova 9
Physical terms
  energy 58
  force 58
  power 58
  pressure 58
  work 58
Piercy, J. E. 171
Pin 1 problem 588
  solution 589
Pink noise 40, 426
Pink noise spectra 426
Pinna acoustic response 27
Pinnae response 267, 268
Pinnae responses 30
Piston impedance function 301, 349
Piston velocity 350
Planck area 49
Planck constant 48
Planck frequency 48
Planck length 42, 48
Planck mass 47
Planck scales 47
Planck system 47
Planck time 48
Planck units 47
Planck, Max 19, 35, 47
Planck’s constant 19
Plane
  complex 77
Plane wave 334
Plane Wave Tube 298
Plane wave tube 291, 294, 296, 297, 299, 304, 371
Plane waves 288, 291, 296, 299, 305, 307
Plot
  ETC 195
  Nyquist 200
  polar envelope time (PET) 201
Plots

  directivity 155
Plug
  RJ-45 604
Polar lunes 156
Polarity
  absolute 576
  loudspeaker 576
polarity 192
Pole 507
Pole and zero analysis 508
Pole-zero diagram 509
Potential acoustic gain (*PAG*) 256, 259
Power
  acoustic 93
  available input 140
  average 118
  compression 427
  electric 117
  electrical 90
  from a voltage source 431
  phantom 340
  radiated 358
  relative speech 270
  response 384
Power factor 119
Power rating 428
Power ratings for amplifiers and loudspeakers
  pink noise 426
Power response 384
Powered loudspeakers 439
Predicting room reflections 450
Pre-emphasis 39
Pressure
  acoustic 347, 351
    far field 354
    near field 354
  sound 178
  static 349
Pressure gradiant 325
Pressure gradient 326
Pressure microphones 325
Pressure Zone Microphone (PZM) 226
Pribram, Karl H. 19, 20
Probability theory 47
Process
  asynchronous 603
  isochronous 603
  synchronous 603
Processing
  dynamics 528
Processor
  virtual sound 581
Psychoacoustic effect 556
Psychoacoustics
  bands 28
  barks 28
  binaural recording 27
  critical band 28
  critical bandwidth 30
  equal loudness contours 30
  equivalent rectagular bandwidth 29
  equivalent rectangular bandwidths (ERBs) 28
  head related transfer functions (HRTFs) 27
  ITE (in the ear—at the eardrum) 27
  motivations 25
  phon level 29
  phons 28
  pinna acoustic response 27
  sones 28, 29
  sound reproduction 25
Pulsating sphere 305
Pupin, Dr. Michael I. 11
Purdue University 15
Putting it all together
  acoustical analysis 579
  analog interconnections 583
  balanced circuits 583
  circuits
    AES3
      connector 599
      encoding 597
      transmission characteristics 598
    cable
      AES3 599
      attenuation 592
      Cat-5 604
      phase velocity 592
      propagation properties 595
    CobraNet™ 603
      interface features 604
    connector
      XLD 602
    connectors
      TSR 588
      XLR 588
    Ethernet 604
    impedance
      characteristic 592
    microphone
      connector
        digital 602
      digital output 601
    pin 1 problem 588
      solution 589
    plug
      RJ-45 604
    process
      asynchronous 603
      isochronous 603
      synchronous 603
    S/PDIF 600
    SCIN removal 589
    skin effect 593
    unbalanced 587
    unbalanced to balanced input 589
  device interconnections 582
  electomagnetic waves
    absorption 585
  electromagnetic waves
    polarization 585
    reflection 585
    skin depth 585
  EMI susceptibility 584
  loudspeaker
    wire loss 582
  loudspeaker management systems 580
  MediaMatrix 581
  virtual sound processor 581

Q

Q 69, 155, 156, 157, 158, 271, 277
  calculating minimum 271
  geometric 156
  by placement 159
  definition 156
  improper 275
  loudspeaker directivity factor 155
  minimum 271
  role in %AL_{CONS} 277
  transform to steradians (sr) 156
Q in %AL_{CONS} 277
Quadratic residue diffusors 248
Quality factor (*Q*) 351
Quality of Service (QoS) 549
Quantization 535, 550
Quantizing 35

R

Radian 75
Radiating surface area 347

Radiation
  half space 353
  impedance 349
Radiation impedance 349
Radiation properties of a loudspeaker 445
Radiation resistance 359
RASTI 278
Rate of decay of reverberant sound energy 218
Ratios 70
  acceptable room ratios 242
  directivity 160
  numbers 54
  power 88, 89
  room 240
  voltage 88
Ray behavior 454
Ray tracing 449, 460
Rayleigh, Lord 15, 35, 369
Reactance
  capacitive 116
  inductive 116
  making visible 136
Real world power generation 434
Realistic room models 457
Reeves, Alex Harley 35
Reflection 176
  focused 273
  specific 274
  specific long, delayed, high level 274
Reflection free zone 244
Reflections per second (*RPS*) 218
Refraction 171, 176
Regeneration 255
Regenerative swelling 256
Release time 531
Renard Series 66
Required electrical power (*REP*) 261
Resistance
  radiation 359
Resistors 113, 125
Resolution
  frequency 198
  time 198
Resonance
  free air 352
  parallel 124
  series 120
Response
  impulse 195
  magnitude 196
  minimum phase 196
  phase 196
  pinnea 207
Reverberant energy 456
Reverberant sound field 451
Reverberant sound level (L_R) 224
Reverberation 456
  excessive 274
Reverberation time
  small room 238
Reverberation time (L_T) 224
Rice, Chester W. 347
Rice, Edwin W. 11
Rigging 4
RIR 456
RJ-45 604
Rodgers, Carolyn “Puddie” 244, 465
Room
  absorption 574
Room absorption 230
Room acoustics 451
Room impulse response (*RIR*) 451
Room measurement nomenclature 608
Room model detail 449
RoomCapture™ 457
Root
  numbers 54
Root mean square 118
RT_{60} 217, 226
  calculating maximum 271
  excessive 274
Rudmose, Dr. Wayne 555
Rudmose, Wayne 13

S

S plane 507
S plane to Z plane mapping 522
S/PDIF 600
Sabine equation 215, 217
Sabine, Wallace Clement 15, 216, 219, 453
Sacks, Oliver 26
Sample and hold 495, 497
Sampling 35
Sampling and recovery 489–495
Sampling angular frequency 493
Sampling Theorem 493–501
Santa Fe Opera 38
Scales
  linear 66
  logarithmic 66
Schotto, Gaspare P. 9
Schroeder, Manfred 215, 248
Schultz, Ted 238
SCIN removal 589
Scott, Herman Hosmer 15
Seals, Dan 575
Seeing Voices 26
Series resonance 120
Setup for measuring the MIV 427
Seven Arts 13
Shannon entropy 41
Shannon Information Theory 35
Shannon Space 36
Shannon, Claude 12, 25, 37, 40, 43, 47
Shannon’s theory 35
SI to US conversion 57
Siemens, Ernst Werner von 10
Signal
  analytic 195
  Nyquist
    trace 196
Signal alignment 153
Signal delay 465
  acoustic astigmatism 470
  acoustic center 470
  acoustic origins 470
  alignment 469
  convergence 469
  delay 467
  directional control device (DCD) 470
  distance 467
  Doak and Bolt delay 469
  effect of combining unaligned loudspeakers in an array 472
  equations 466
  Fay-Hall effect 468
  finding the velocity of sound with an analyzer 467
  frequency response of unaligned and synchronized horns 473
  Gold Line TEF analyzer 468
  Haas effect 468
  Henry effect 468
  loudspeaker missynchronization 470
  natural signal delay 469
  null frequency interval *(NFI)* 470
  polar response 471
  setting signal delay 469
  signal alignment 469
  signal synchronization 469

  synchronization & alignment of arrays 469
  time 467
  two loudspeakers misaligned 471
  under-balcony loudspeaker delay 469
  useful 466
  velocity 467
  velocity equation 467
  why delay 465
Signal processing 510
  ADC 500
  aliasing 494
  analog to digital conversion 497
  attack time 530
  Bessel polynomials 510
  binary codes 498
  biquad 526
  Butterworth polynomials 510
  compression 529
  DAC 499
  delta sigma modulation 501
  dynamics processing 528
  equalization 514
    global 517
    local 517
  Euler’s theorem 481
  expansion 529
  filter
    FIR 522
    recursive 524
  Fourier trigonometric series 478
  impulse sampling 489
  Laplace transform 503, 520
  limiting 530
  minimum phase 513
  non-linear dynamics processing 529
  non-minimum phase 513
  number base 498
  Parseval’s theorem 485
  phase
    linear 522
  pole 507
  pole and zero analysis 508
  pole-zero diagram 509
  release time 531
  S plane 507
  S plane to Z plane mapping 522
  sample and hold 495, 497
  Sampling Theorem 493–501
  spectra 477
  spectral density 486
  spectrum of a periodic impulse sampled signal 489
  staircase time domain behavior 496
  system theory 502
  table of function transform pairs 503
  table of operations transform pairs 504
  transfer function 506
  transfer function surface 507
  uncertainly principle, classical 487
  unit impulse 487
Signal-to-noise ratio (*SNR*) 225, 253, 271
Sine wave
  steady state 192
  swept 192
Single-mode fiber 545
Singularity problem 42
Sketchup™ 460
Skin effect 593
Small room acoustics
  acceptable ratios 242
  Acorn Studios 248
  acoustical parameters 238
  axial modes 240, 241
  damped and undamped modes 239
  dead end 243
  diffusion 244, 245
  geometry 240
  Haas effect 243
  initial signal delay gap (ISD) 240, 241
  ISD 240, 241
  LEDE 240
    concepts 242
    design 248
  live end 243
  Live End Dead End 240
  Master Sound Astoria Recording Studio 243
  modal decay rates 239
  modes 239
  Non-statistical spaces 237
  oblique mode 241
  quadratic residue diffusors 248
  ratios 240
  reflection free zone 244
  resonances 239
  reverberation time 238
  tangential mode 241
Smolin, Lee 22
Snow, William B. 12, 226, 558, 559, 572
SNR 253, 268, 274, 320
Sones 28, 29
Sonex 246
Sony/Philips Digital Interface (S/PDIF) 540
Sound
  effect of altitude on the velocity of sound in air 175
  intensity 178
  pressure 178
  temperature effect on speed 466
  velocity 173
Sound and our brain
  axon chips 19, 22
  conscious electromagnetic information field (CEMI) 21
  digital voice recognition 22
  Fourier transform 19
  human brain 19
  neuron chips 19, 22
  synapse chips 19, 22
  transcranial magnetic stimulation 21
Sound behavior 454
Sound energy density 178
Sound intensity 178
Sound intensity level (L_I) 220
Sound level meter (SLM) 315
Sound power level (L_W) 220
Sound pressure 178
Sound pressure level 178
  axial 363
Sound pressure level (L_P) 220
Sound radiation patterns 259
Sound reproduction 25
Sound System Engineering 15, 443
Sound system equalization
  absolute polarity 576
  acoustic feedback
    analysis 556–558
    decay 557
    growth 556
    negative growth and decay 558
  active band-pass filter 564
  band-boost filter 562, 567
  bandpass filter 562, 567
  band-rejection filter 562, 567
  bridged-T filter 566
  broad-band combining-type

  band-rejection filter 565
  criteria for band-rejection filters 562
  diaphragmatic absorbers 574
  direct sound level (L_D) 555
  don’t equalize for hearing loss 575
  early practitioners 559
  effects of surfaces on feedback 575
  envelope time curve 568–569
    magnitude and phase 569
  equalizing for playback 573
  ETC 568
  feedback
    defined 556
  feedback is a single frequency 570
  feedback path 556
  feedback stability margin 558
  FFT 573
  filter parameters 562
  filter transfer characteristics 565
  forward transfer function 556
  gain ratio 557
  GenRad 559
  group delay 555
  Hewlett Packard 559
  high-quality wave analyzer 573
  house curves 561, 574
  improper use of real time analysis 574
  loudspeaker polarity 576
  microphone polarity 575
  minimum phase loudspeaker 567
  minimum phase response 562
  minimum-phase filters 567
  Mother Natures way 572
  multi-loudspeaker arrays 570
  multiplicity of sound fields 570–571
  narrow-band filter 564
  negative acoustic feedback 558
  noise control 555
  Noise Criteria curves 560
  Nyquist display 569
  oscillatory feedback 556
  passive equalizer 564
  pink noise 561
  positive feedback 556, 558
  properly isolated 572
  properly shock mounted 572
  proximity modes 575
  psychoacoustic effect 556
  real time regenerative response method 573
  real-time analyzer 559, 573, 576
  real-time regenerative-response 572
  regenerative response 572
    tuning 573
  room absorption 574
  RTA application 561
  sound field 570
  sources of feedback 569
  steady state gain ratio 558
  steady-state-feedback 572
  sweep oscillators 573
  TEF analysis 567, 576
  TEF analyzer 573, 576
  time dependency 570
  transient nature of acoustic feedback 556
  unnatural distortions 572
  white noise 561
Source
  impulse 192
Source impedance 323
Space
  non-statistical 237
Specific acoustic impedance 293, 296, 300, 305, 306
Spectra 477
Spectral density 486
Specular reflections 460
Speech intelligibility 269
  %AL_{CONS} 268, 269, 272, 276, 277
  %AL_{CONS} variables 277
  acoustic environment 267
  articulation 270
  articulation index 271
  articulation losses of consonants in speech 268
  calculating %AL_{CONS} 269
  calculating the minimum 271
  calculation of the maximum RT_{60} 271
  calculations 271
  causes of reduced intelligibility 274
  changing STI measurements to %AL_{CONS} 274
  choosing the correct %AL_{CONS} equation 277
  comb filter 276
  converting STI measurements to %AL_{CONS} 274
  detecting reflections without analyzers 278
  distance from the source 274, 275
  effect of *SNR* on %AL_{CONS} 271
  electronic systems 267
  envelope time curve 269, 274
  ETC 269, 274
  evaluating speech intelligibility measurements 273
  excessive reverberation 274
  excessive RT_{60} 274
  factors affecting %AL_{CONS} 272
  focused reflections 273
  high density overhead distribution 276
  Hopkins-Stryker equation 277
  improper coverage angles 275
  improper *Q* 275
  Indiana University 267
  Intelligibility Workshop 267
  Intelligibility Workshop 1986 278
  key parameters affecting speech intelligibility 268
  lack of synchronization 274
  loudspeaker misalignment 274
  loudspeaker system 267
  Ma 277
  Maxfield’s equation 270
  maximum RT_{60} 271
  microphone 267
  minimum *Q* 271
  misalignment 275
  misequalization 274, 275
  modified rhyme tests 278
  modular transfer function 268
  MRT tapes by Dynastat 278
  MTF 268
  non-acoustic articulation problems 275
  pew back loudspeaker systems 276
  pinnae response 267, 268
  polar responses before and after alignment 276
  poor quality devices 274
  poor *SNR* 274
  RASTI 278
  receiver-listener 267
  relationship between acoustic gain and %AL_{CONS} 278
  relationship between Q_{MIN} and

  $D_{2\ (MAX)}$ 276
  relative speech power 270
  role of 277
  SNR 268, 271, 274
  Sonex 278
  specific long, delayed, high-level reflections 274
  specific reflections 274
  speech power 270
  Speech Transmission Index 273
  STI 273
  subjective word tests 278
  synchronization 274
  talker 267
  TEF analyzers 269
  TEF measurements 278
  usable percentages 269
Speech Transmission Index 273
Sphere 158
  divided into polar lunes 163
  divided into zones 163
  hemisphere 158
  one-eighth sphere 158
  one-quarter 158
  spherical segments 158
Spherical loudspeaker data 443
Spherical wave 444
Spherical waves 306
Staircase time domain behavior 496
Standard data formats 540
Standard Handbook for Electrical Engineers 5
Standing wave 299, 303
Static air pressure 349
Steinberg 12
Steinberg, J. C. 559
Steinmetz, Charles Proteus 11, 78
Steradians, sr 156
STI 273
Stochastic process 40
Stockham, Thomas 38
Strutt, M. J. O. 216
Stuart, J. Robert 41
Studer 38
SU 87
Sullivan, Anne 26
Surface acceleration 347
Surface displacement 347
Surface velocity 347
Susceptance 124
Sutherland, L. C. 171
Symbols 605
Symbols and abbreviations 605
Synapse chips 19, 22
Syn-Aud-Con Newsletters 15
Synchronization 537, 539
  lack of 274
Synchronous process 603
Synergetic Audio Concepts 13
System
  criteria 555
  electrical gain 142
  electronic 267
  loudspeaker 267
System gain structure 439
System theory 502
Systems
  digital 520–528
  playback 108

T

*T*30 456
Table of function transform pairs 503
Table of operations transform pairs 504
Tables
  math 72
Tangential modes 240, 241
Taylor, Herb 13
TEF 193, 240
  analysis 567
  analyzer 269, 573, 576
  displays 196
  processing signals 194
Tektronics 14
Telephone circuits 87
Temperature
  effect on the speed of sound 466
Tesla, Nikola 10
Test
  eye pattern 536
Text into binary, octal, hexadecimal 46
The Mathematical Theory of Communication 43
The Physics of Information 41
The Theory of Probability 47
Thermal noise 317
Thermodynamic entropy 41
Things amplifiers hate
  2 ohm loads 435
  low crest factor signals 435
Thomson, William 113
Thuras, Albert L 12
Time 465, 467
  clock time 465
  dependency 570
  resolution 198
Time Delay Spectrometry (TDS) 13
Todrank, Robert 248
Tone meister 28
Tononi, Giulio 20
Total power (L_W) 221
Transcranial magnetic stimulation (TMS) 21
Transfer function 379, 506
Transfer function surface 507
Transport of digital audio data 537
Tronchin, Lamberto 9
Tube
  plane wave 371
Tukey, John W. 35, 44
Tuller, Betty 40
Turing, Alan 47
Two port devices 146

U

Unbalanced circuits 587
Unbalanced to balanced input 589
Uncertainty principle, classical 487
Unit impulse 487
Universal gas constant 284
US to SI conversion 57

V

Velocity 80, 467
  phase 592
  temperature dependant 174
VI 12
Ville, J. 43
Virtual Local Area Network (VLAN) 548, 549
Virtual sound processor 581
Volta, Alessandro 9
Voltage 88
  amplitude 88
  combining 99
  ratio 88
Volume velocity 347
Von Guericke, Otto 9

W

Wahrenbrock, Kenneth 333
Warner Bros. 13
Watt, James 10
Wave behavior 454

Waveform
  analog 535
Wavelength 175, 287, 292, 295, 303, 310, 350
  frequency 175
  typical 175
Waves
  electromagnetic
    absorption 585
    polarization 585
    reflection 585
    skin depth 585
Weber, Wilhelm Eduard 10
Webster, A. G. 369
Wente, Edward C. 12
Western Electric 13
Wheeler, John Archibald 38, 47
White noise 40
Wilczek, Frank 42
Wilson, Fred 13
Wire
  loss 582
Wire-frame model 458
Word clock 539

X

XLD 602

Z

Z transform 520
Zone
  reflection free 244